

FIGURE 1

GGCTGAGGGGAGGCCCGGAGCCTTTCTGGGGCCTGGGGGATCCTCTTGCACTGGTGGGTGGA
GAGAAGCGCCTGCAGCCAACCAGGGTCAGGCTGTGCTCACAGTTTCCTCTGGCGGCATGTAA
AGGCTCCACAAAGGAGTTGGGAGTTCAAATGAGGCTGCTGCGGACGGCCTGAGGATGGACCC
CAAGCCCTGGACCTGCCGAGCGTGGCACTGAGGCAGCGGCTGACGCTACTGTGAGGGAAAGA
AGGTTGTGAGCAGCCCCGCAGGACCCCTGGCCAGCCCTGGCCCCAGCCTCTGCCGGAGCCCT
CTGTGGAGGCAGAGCCAGTGGAGCCAGTGGAGGCAGGGCTGCTTGGCAGCCACCGGCCTGCA
ACTCAGGAACCCCTCCAGAGGCCATGGACAGGCTGCCCCGCTGACGGCCAGGGTGAAGCATG
TGAGGAGCCGCCCCGGAGCCAAGCAGGAGGGAAGAGGCTTTCATAGATTCTATTACAAAGA
ATAACCACCATTTTGAAGGACC**ATG**AGGCCACTGTGCGTGACATGCTGGTGGCTCGGACTG
CTGGCTGCCATGGGAGCTGTTGCAGGCCAGGAGGACGGTTTTGAGGGCACTGAGGAGGGCTC
GCCAAGAGAGTTTCAATTTACCTAAACAGGTACAAGCGGGCGGGCGAGTCCCAGGACAAGTGCA
CCTACACCTTCATTGTGCCCCAGCAGCGGGTCACGGGTGCCATCTGCGTCAACTCCAAGGAG
CCTGAGGTGCTTCTGGAGAACCGAGTGCATAAGCAGGAGCTAGAGCTGCTCAACAATGAGCT
GCTCAAGCAGAAGCGGCAGATCGAGACGCTGCAGCAGCTGGTGGAGGTGGACGGCGGCATTG
TGAGCGAGGTGAAGCTGCTGCGCAAGGAGAGCCGCAACATGAACTCGCGGGTCACGCAGCTC
TACATGCAGCTCCTGCACGAGATCATCCGCAAGCGGGACAACGCGTTGGAGCTCTCCAGCT
GGAGAACAGGATCCTGAACCAGACAGCCGACATGCTGCAGCTGGCCAGCAAGTACAAGGACC
TGGAGCACAAGTACCAGCACCTGGCCACACTGGCCACACAACCAATCAGAGATCATCGCGCAG
CTTGAGGAGCACTGCCAGAGGGTGCCCTCGGCCAGGCCCGTCCCCCAGCCACCCCCCGCTGC
CCCGCCCCGGGTCTACCAACCACCCACCTACAACCGCATCATCAACCAGATCTCTACCAACG
AGATCCAGAGTGACCAGAACCTGAAGGTGCTGCCACCCCTCTGCCCACTATGCCCACTCTC
ACCAGCCTCCCATCTTCCACCGACAAGCCGTCGGGCCCATGGAGAGACTGCCTGCAGGCCCT
GGAGGATGGCCACGACACCAGCTCCATCTACCTGGTGAAGCCGGAGAACACCAACCGCCTCA
TGCAGGTGTGGTGCGACCAGAGACACGACCCCGGGGGCTGGACCGTCATCCAGAGACGCCTG
GATGGCTCTGTAACTTCTTTCAGGAACCTGGGAGACGTACAAGCAAGGGTTTGGGAACATTGA
CGGCGAATACTGGCTGGGCCTGGAGAACATTTACTGGCTGACGAACCAAGGCAACTACAAAC
TCCTGGTGACCATGGAGGACTGGTCCGGCCGCAAAGTCTTTGCAGAATACGCCAGTTTCCGC
CTGGAACCTGAGAGCGAGTATTATAAGCTGCGGCTGGGGCGCTACCATGGCAATGCGGGTGA
CTCCTTTACATGGCACAACGGCAAGCAGTTCACCACCCTGGACAGAGATCATGATGTCTACA
CAGGAAACTGTGCCCACTACCAGAAGGGAGGCTGGTGGTATAACGCCTGTGCCCACTCCAAC
CTCAACGGGGTCTGGTACCGCGGGGGCCATTACCGGAGCCGCTACCAGGACGGAGTCTACTG
GGCTGAGTTCCGAGGAGGCTCTTACTCACTCAAGAAAGTGGTGATGATGATCCGACCGAACC
CCAACACCTTCCACT**TAAG**CCAGCTCCCCCTCCTGACCTCTCGTGGCCATTGCCAGGAGCCCA
CCCTGGTCACGCTGGCCACAGCACAAAGAACAACCTCCTACCAAGTTCATCCTGAGGCTGGGA
GGACCGGGATGCTGGATTCTGTTTTCCGAAGTCACTGCAGCGGATGATGGAACTGAATCGAT
ACGGTGTTTTCTGTCCCTCCTACTTTCCTTCACACCAGACAGCCCCTCATGTCTCCAGGACA
GGACAGGACTACAGACAACCTTTCTTTAAATAAATTAAGTCTCTACAATAAAAAAAA

2/75

FIGURE 2

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA22779

><subunit 1 of 1, 493 aa, 1 stop

><MW: 57104, pI: 7.67, NX(S/T): 2

MRPLCVTCWWLGLLAAMGAVAGQEDGFEGTEEGSPREFIYLNRYKRAGESQDKCTYTFIVPQ
 QRVGTGAICVNSKEPEVLLLENRVHKQELELLNNELLKQKRQIETLQQLVEVDGGIVSEVKLLR
 KESRNMNSRVTQLYMQLLHEIIRKRDNALELSQLENRIILNQTADMLQLASKYKDLEHKYQHL
 ATLAHNQSEIIAQLLEHCQRVPSARVPVQPPPAAPPRVYQPPTYNRIINQISTNEIQSDQNL
 KVLPPPLPTMPTLTSLPSSTDKPSGPWRDCLQALEDGHDTSIYLVKPENTNRLMQVWCDQR
 HDPGGWTVIQRRLDGSVNFNRNWETKQGGFNIDGEYWLGLENIYWLTNQGNKLLVTMEDW
 SGRKVF AEYASFRLEPESEYYKLRLGRYHGNAGDSFTWHNGKQFTTLDRDHDVYTGNC AHYQ
 KGGWWYNACAHSNLNGVWYRGGHYRSRYQDGVYWAEFRGGSYSLKKVMMIRPNPNTFH

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 164-168, 192-196

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 124-128

Tyrosine kinase phosphorylation sites.

amino acids 177-184, 385-393, 385-394, 461-468

N-myristoylation sites.

amino acids 12-18, 18-24, 22-28, 29-35, 114-120, 341-347,
 465-471, 473-479

Amidation site.

amino acids 373-377

Fibrinogen beta and gamma chains C-terminal domain signature.

amino acids 438-451

Fibrinogen beta and gamma chains C-terminal domain proteins.

amino acids 305-343, 365-402, 411-424, 428-458

Trehalase proteins.

amino acids 275-292

3/75

FIGURE 3

CCCACGCGTCCGGCGCCGTGGCCTCGCGTCCATCTTTGCCGTTCTCTCGGACCTGTCACAAA
GGAGTCGCGCCGCGCCGCGCCGCCCCCTCCCTCCGGTGGGCCCCGGGAGGTAGAGAAAGTCAGT
GCCACAGCCCGACCGCGCTGCTCTGAGCCCTGGGCACGCGGAACGGGAGGGAGTCTGAGGGT
TGGGGACGTCTGTGAGGGAGGGGAACAGCCGCTCGAGCCTGGGGCGGGCGGACCGGACTGGG
GCCGGGGTAGGCTCTGGAAAGGGCCCCGGGAGAGAGGTGGCGTTGGTCAGAACCTGAGAAACA
GCCGAGAGGTTTTCCACCGAGGCCCCGCGCTTGAGGGATCTGAAGAGGTTCCCTAGAAGAGGGT
GTTCCCTCTTTCGGGGGTCCTCACCAGAAGAGGTTCTTGGGGGTCGCCCTTCTGAGGAGGCT
GCGGCTAACAGGGCCCCAGAACTGCCATTGGATGTCCAGAATCCCCTGTAGTTGATAATGTTG
GGAATAAGCTCTGCAACTTTCTTTGGCATTTCAGTTGTTAAAAACAAATAGGATGCAAATTCC
TCAACTCCAGGTTATGAAAACAGTACTTGGAAAACAGAAAACCTAAATGATCGTCTTTG
GTTGGGCCGTGTTCTTAGCGAGCAGAAGCCTTGGCCAGGGTCTGTTGTTGACTCTCGAAGAG
CACATAGCCCCTTCCTAGGGACTGGAGGTGCCGCTACTACCATGGGTAATTCCTGTATCTG
CCGAGATGACAGTGGAACAGATGACAGTGTTGACACCCAACAGCAACAGGCCGAGAACAGTG
CAGTACCCACTGCTGACACAAGGAGCCAACCACGGGACCCTGTTTCGGCCACCAAGGAGGGGC
CGAGGACCTCATGAGCCAAGGAGAAAGAAACAAAATGTGGATGGGCTAGTGTGAGACACACT
GGCAGTAATACGGACTCTTGTAGATAAGTAAAGTATCTGACTCACGGTCACCTCCAGTGGAAT
GAAAAGTGTTCTGCCCCGGAACCATGACTTTAGGACTCCTTCAGTTCCCTTTAGGACATACTCG
CCAAGCCTTGTGCTCACAGGGCAAAGGAGAATATTTTAATGCTCCGCTGATGGCAGAGTAAA
TGATAAGATTTGATGTTTTTGCTTGCTGTCATCTACTTTGTCTGGAAATGTCTAAATGTTTC
TGTAGCAGAAAACACGATAAAGCTATGATCTTTATTAGAG

...
...
...

4/75

FIGURE 4

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26846

<subunit 1 of 1, 117 aa, 1 stop

<MW: 12692, pI: 7.50, NX(S/T): 0

MIVFGWAVFLASRSLGQGLLLTLEEHIAHFLGTGGAATTMGNSCICRDDSGTDDSVDTQQQQ

AENSAVPTADTRSQPRDPVRPPRRGRGPHEPRRKKQNV DGLVLDTLAVIRTLVDK

Important features:

Signal peptide:

amino acids 1-16

N-myristoylation sites.

amino acids 18-24, 32-38, 34-40, 35-41, 51-57

FIGURE 5

CCCACGCGTCCGCGCAGTCGCGCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGG
CCCGGGGCTGTGGCGTCGACTCCGACCCAGGCAGCCAGCAGCCCGCGCGGGAGCCGGACCGC
CGCCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA
GCCCCGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGGAGACAGCGGACAAGCAG
CGGAGGAGAAGGAGGAGGAGGCGAACCCAGAGAGGGGCGAGCAAAAGAAGCGGTGGTGGTGGG
CGTCGTGGCC**ATG**GCGGCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCG
AGCGCGAGAAATCCAACGCCTGCAAGTGTGTCAGCAGCCCCAGCAAAGGCAAGACCAGCTGC
GACAAAAACAAGTTAAATGTCTTTTCCCGGGTCAAACCTCTTCGGCTCCAAGAAGAGGCGCAG
AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC
ACTTGCAGCTGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACT
CTGTTTAACCTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT
GTA

amino acids 78-94, 102-153

7/75

FIGURE 7

ATGGCCGCGGCCATCGCTAGCGGCTTGATCCGCCAGAAGCGGCAGGCGCGGGAGCAGCACTG
GGACCGGCCGTCTGCCAGCAGGAGGCGGAGCAGCCCCAGCAAGAACCGCGGGGCTCTGCAACG
GCAACCTGGTGGATATCTTCTCCAAAGTGCGCATCTTCGGCCTCAAGAAGCGCAGGTTGCGG
CGCCAAGATCCCCAGCTCAAGGGTATAGTGACCAGGTTATATTGCAGGCAAGGCTACTACTT
GCAAATGCACCCCGATGGAGCTCTCGATGGAACCAAGGATGACAGCACTAATTCTTACTCTCT
TCAACCTCATACCAGTGGGACTACGTGTTGTTGCCATCCAGGGAGTGAAAACAGGGTTGTAT
ATAGCCATGAATGGAGAAGGTTACCTCTACCCATCAGAACTTTTTTACCCCTGAATGCAAGTT
TAAAGAATCTGTTTTTTGAAAATTATTATGTAATCTACTCATCCATGTTGTACAGACAACAGG
AATCTGGTAGAGCCTGGTTTTTTGGGATTAAATAAGGAAGGGCAAGCTATGAAAGGGAACAGA
GTAAAGAAAACCAAACCAGCAGCTCATTTTCTACCCAAGCCATTGGAAGTTGCCATGTACCG
AGAACCATCTTTGCATGATGTTGGGGAAACGGTCCCGAAGCCTGGGGTGACGCCAAGTAAAA
GCACAAGTGCGTCTGCAATAATGAATGGAGGCAAACCAGTCAACAAGAGTAAGACAACAT**TAG**

8/75

FIGURE 8

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA28503

><subunit 1 of 1, 247 aa, 1 stop

><MW: 27702, pI: 10.36, NX(S/T): 2

MAAAIASGLIRQKRQAREQHWDRPSASRRRSPSKNRGLCNGNLVDIFSKVRIFFGLKKRRLR
RQDPQLKGIVTRLYCRQGYLQMHPDGALDGTKDDSTNSTLFNLIIPVGLRVVAIQGVKTGLY
IAMNGEGYLYPSELFTEPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNR
VKKTKPAAHFLPKPLEVAMYREPSLHDVGETVVPKPGVTPSKSTSASAIMNGGKPVNKSSTT

N-glycosylation site.

amino acids 100-104, 242-246

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 28-32, 29-33

Tyrosine kinase phosphorylation site.

amino acids 199-207

N-myristoylation site.

amino acids 38-44, 89-95, 118-124, 122-128, 222-228

HBGF/FGF family proteins.

amino acids 104-155, 171-198

FIGURE 9

CTCGCAGCCGAGCGCGGCCGCGGGAAGGGCTCTCCTTCCAGCGCCGAGCACTGGGCCCTGGCA
GACGCCCCAAGATTGTTGTGAGGAGTCTAGCCAGTTGGTGAGCGCTGTAATCTGAACCAGCT
GTGTCCAGACTGAGGCCCATTTGCATTGTTTAAACATACTTAGAAAATGAAGTGTTCATTTT
TAACATTCTCTCTCCAATTGGTTTAAATGCTGAATTACTGAAGAGGGCTAAGCAAAACCAGGT
GCTTGCGCTGAGGGCTCTGCAGTGGCTGGGAGGACCCCGCGCTCTCCCCGTGTCTCTCCA
CGACTCGCTCGGCCCTCTGGAATAAAACACCCGCGAGCCCCGAGGGCCCAGAGGAGGCCGA
CGTGCCCGAGCTCCTCCGGGGGTCCCGCCCGCGAGCTTTCTTCTCGCCTTCGCATCTCTCTCC
TCGCGCGTCTTGGACATGCCAGGAATAAAAAGGATACTCACTGTTACCATTCTGGCTCTCTG
TCTTCCAAGCCCTGGGAATGCACAGGCACAGTGCACGAATGGCTTTGACCTGGATCGCCAGT
CAGGACAGTGTTTAGATATTGATGAATGCCGAACCATCCCCGAGGCCTGCCGAGGAGACATG
ATGTGTGTTAACCAAAATGGCGGGTATTTATGCATTCCCCGACAAAACCTGTGTATCGAGG
GCCCTACTCGAACCCCTACTCGACCCCTACTCAGGTCCGTACCCAGCAGCTGCCCCACCAC
TCTCAGCTCCAACTATCCCACGATCTCCAGGCCTCTTATATGCCGCTTTGGATAACCAGATG
GATGAAAGCAACCAATGTGTGGATGTGGACGAGTGTGCAACAGATTCCCACCAGTGCAACCC
CACCCAGATCTGCATCAATACTGAAGGCGGGTACACCTGCTCCTGCACCGACGGATATTGGC
TTCTGGAAGGCCAGTGCTTAGACATTGATGAATGTGCTATGGTTACTGCCAGCAGCTCTGT
GCGAATGTTCTTGATCCTATTCTTGTACATGCAACCTGGTTTTACCCTCAATGAGGATGG
AAGGTCTTGCCAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTGCAAACCTGCGTCA
ACACCTACGGCTCTCTCATCTGCCGCTGTGACCCAGGATATGAACCTGAGGAAGATGGCGTT
CATTGCAGTGATATGGACGAGTGCAGCTTCTCTGAGTTCCTCTGCCAACATGAGTGTGTGAA
CCAGCCCCGGCACATACTTCTGCTCCTGCCCTCCAGGCTACATCCTGCTGGATGACAACCGAA
GCTGCCAAGACATCAACGAATGTGAGCACAGGAACCACACGTGCAACCTGCAGCAGACGTGC
TACAATTTACAAGGGGGCTTCAAATGCATCGACCCCATCCGCTGTGAGGAGCCTTATCTGAG
GATCAGTGATAACCGCTGTATGTGTCTGCTGAGAACCCTGGCTGCAGAGACCAGCCCTTTA
CCATCTTGTACCGGGACATGGACGTGGTGTGTCAGGACGCTCCGTTCCCGCTGACATCTTCCAA
ATGCAAGCCACGACCCGCTACCCTGGGGCCTATTACATTTTCCAGATCAAATCTGGGAATGA
GGGCAGAGAATTTTACATGCGGGCAAACGGGGCCCCATCAGTGCCACCCCTGGTGATGACACGCC
CCATCAAAGGGCCCCGGGAAATCCAGCTGGACTTGGAAATGATCACTGTCAACACTGTCATC
AACTTCAGAGGCAGCTCCGTGATCCGACTGCGGATATATGTGTGCGCAGTACCCATTCTGAGC
CTCGGGCTGGAGCCTCCGACGCTGCCTCTCATTGGCACCAAGGGACAGGAGAAGAGAGGAAA
TAACAGAGAGAATGAGAGCGACACAGACGTTAGGCATTTCTGCTGAACGTTTCCCCGAAGA
GTCAGCCCCGACTTCCTGACTCTCACCTGTACTATTGCAGACCTGTACCCCTGCAGGACTTG
CCACCCCCAGTTCCCTATGACACAGTTATCAAAAAGTATTATCATTGCTCCCCTGATAGAAGA
TTGTTGGTGAATTTTCAAGGCCTTCAGTTTATTTCCACTATTTTCAAAGAAAATAGATTAGG
TTTGCGGGGTCTGAGTCTATGTTCAAAGACTGTGAACAGCTTGCTGTCACTTCTTCACCTC
TTCCACTCCTTCTCTCACTGTGTTACTGCTTTGCAAAGACCCGGGAGCTGGCGGGGAACCT
GGGAGTAGCTAGTTTGCTTTTTCGTACACAGAGAAGGCTATGTAAACAAACCACAGCAGGA
TCGAAGGGTTTTTAGAGAATGTGTTTCAAACCATGCCTGGTATTTTCAACCATAAAAGAAG
TTTCAGTTGTCCTTAAATTTGTATAACGGTTTAAATCTGTCTTGTTCATTTTGAGTATTTTT
AAAAAATATGTCGTAGAATTCCTTCGAAAGGCCTTCAGACACATGCTATGTTCTGTCTTCCC
AAACCCAGTCTCCTCTCCATTTTAGCCAGTGTTTTCTTTGAGGACCCCTTAATCTTGCTTT
CTTTAGAATTTTTTACCCAATTGGATTGGAATGCAGAGGTCTCCAACTGATTAAATATTTGA
AGAGA

FIGURE 10

MPGIKRILTVTILALCLPSPGNAQAQCTNGFDLDRQSGQCLDIDECRTIPEACRGDMMCVNQ
 NGGYLCIPRTNPVYRGPYSNPYSTPYSGPYPAAPPLSAPNYPTISRPLICRFGYQMDESNQ
 CVDVDECATDSHQCNPTQICINTEGGYTCSCTDGYWLLEGQCLDIDECRYGYCQQLCANVPG
 SYSCTCNPGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSLLICRCDPGYELEEDGVHCSDM
 DECSFSEFLCQHECVNQPGTYFCSCPPGYILLDDNRSCQDINECEHRNHTCNLQQTCYNLQG
 GFKCIDPIRCEEPYLRISDNRCMCPAENPGCRDQPFITILYRDMDVVSGRSVPADIFQMQATT
 RYPGAYYIFQIKSGNEGREFYMRQTGPISATLVMTRPIKGPRIQLDLEMITVNTVINFRGS
 SVIRLRIYVSQYPF

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 283-287, 296-300

N-myristoylation sites.

amino acids 21-27, 64-70, 149-155, 186-192, 226-232, 242-248,
 267-273, 310-316

Aspartic acid and asparagine hydroxylation sites.

amino acids 144-156, 181-193, 262-274

Cell attachment sequence.

amino acids 54-57

Calcium-binding EGF-like.

amino acids 131-166, 172-205, 211-245, 251-286

FIGURE 11

CAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCTCGGGGATCCTCTAGAGATCCCTC
GACCTCGACCCACGCGTCCGAACACAGGTCCTTGTTGCTGCAGAGAAGCAGTTGTTTTGCTG
GAAGGAGGGAGTGC GCGGGCTGCCCCGGGCTCCTCCCTGCCGCTCCTCTCAGTGATGGTT
CCAGGCACCCTGTCTGGGGCAGGGAGGGCACAGGCCTGCACATCGAAGGTGGGGTGGGACCA
GGCTGCCCCCTCGCCCCAGCATCCAAGTCCTCCCTTGGGCGCCCCGTGGCCCTGCAGACTCTCA
GGGCTAAGGTCCTCTGTTGCTTTTTTGGTTCCACCTTAGAAGAGGCTCCGCTTGACTAAGAGT
AGCTTGAAGGAGGCACC**ATG**CAGGAGCTGCATCTGCTCTGGTGGGCGCTTCTCCTGGGCCTG
GCTCAGGCCTGCCCTGAGCCCTGCGACTGTGGGGAAAAGTATGGCTTCCAGATCGCCGACTG
TGCCTACCGCGACCTAGAATCCGTGCCGCTGGCTTCCCGGCCAATGTGACTACACTGAGCC
TGTGAGCCAACCGGCTGCCAGGCTTGCCGGAGGGTGCCTTCAGGGAGGTGCCCTGCTGCAG
TCGCTGTGGCTGGCACACAATGAGATCCGCACGGTGGCCGCCGGAGCCCTGGCCTCTCTGAG
CCATCTCAAGAGCCTGGACCTCAGCCACAATCTCATCTCTGACTTTGCCTGGAGCGACCTGC
ACAACCTCAGTGCCCTCCAATTGCTCAAGATGGACAGCAACGAGCTGACCTTCATCCCCCGC
GACGCCTTCCGCAGCCTCCGTGCTCTGCGCTCGCTGCAACTCAACCACAACCGCTTGCACAC
ATTGGCCGAGGGCACCTTCACCCCGCTCACCGCGCTGTCCACCTGCAGATCAACGAGAACC
CCTTCGACTGCACCTGCGGCATCGTGTGGCTCAAGACATGGGCCCTGACCACGGCCGTGTCC
ATCCCGGAGCAGGACAACATCGCCTGCACCTCACCCCATGTGCTCAAGGGTACACCGCTGAG
CCGCCTGCCGCCACTGCCATGCTCGGCGCCCTCAGTGCAGCTCAGCTACCAACCCAGCCAGG
ATGGTGCCGAGCTGCGGCCTGGTTTTGTGCTGGCACTGCACTGTGATGTGGACGGGCAGCCG
GCCCCCTCAGCTTCACTGGCACATCCAGATACCCAGTGGCATTGTGGAGATCACCAGCCCCAA
CGTGGGCACTGATGGGCGTGCCCTGCCTGGCACCCCTGTGGCCAGCTCCCAGCCGCGCTTCC
AGGCCTTTGCCAATGGCAGCCTGCTTATCCCCGACTTTGGCAAGCTGGAGGAAGGCACCTAC
AGCTGCCTGGCCACCAATGAGCTGGGCAGTGTGAGAGCTCAGTGGACGTGGCACTGGCCAC
GCCCCGTGAGGGTGGTGAGGACACACTGGGGCGCAGGTTCCATGGCAAAGCGGTTGAGGGAA
AGGGCTGCTATACGGTTGACAACGAGGTGCAGCCATCAGGGCCGGAGGACAATGTGGTCATC
ATCTACCTCAGCCGTGCTGGGAACCCTGAGGCTGCAGTCGACAGAGGGGTCCCTGGGCAGCT
GCCCCCAGGCCTGCTCCTGCTGGGCCAAAGCCTCCTCCTCTTCTTCTTCTCCTCACCTCCTTCT**T**
AGCCCCACCCAGGGCTTCCCTAACTCCTCCCCTTGCCCCCTACCAATGCCCCTTTAAGTGCTG
CAGGGGTCTGGGGTTGGCAACTCCTGAGGCCTGCATGGGTGACTTCACATTTTCTTACCTCT
CCTTCTAATCTCTTCTAGAGCACCTGCTATCCCCAACTTCTAGACCTGCTCCAAACTAGTGA
CTAGGATAGAATTTGATCCCCTAACCTCACTGTCTGCGGTGCTCATTGCTGCTAACAGCATTG
CCTGTGCTCTCCTCTCAGGGGCAGCATGCTAACGGGGCGACGTCCTAATCCAACCTGGGAGAA
GCCTCAGTGGTGGAATTCCAGGCACTGTGACTGTCAAGCTGGCAAGGGCCAGGATTGGGGGA
ATGGAGCTGGGGCTTAGCTGGGAGGTGGTCTGAAGCAGACAGGGAATGGGAGAGGAGGATGG
GAAGTAGACAGTGGCTGGTATGGCTCTGAGGCTCCCTGGGGCCTGCTCAAGCTCCTCCTGCT
CCTTGCTGTTTTCTGATGATTTGGGGGCTTGGGAGTCCCTTTGTCTCATCTGAGACTGAAA
TGTGGGGATCCAGGATGGCCTTCCTTCCTCTTACCCTTCCTCCCTCAGCCTGCAACCTCTAT
CCTGGAACCTGTCTCCCTTTCTCCCCAACTATGCATCTGTTGTCTGCTCCTCTGCAAAGGC
CAGCCAGCTTGGGAGCAGCAGAGAAATAAACAGCATTTCTGATGCCAAAAAAAAAAAAAAAAA
AAGGGCGGCCGCGACTCTAGAGTCGACCT

FIGURE 12

MQELHLLWWALLLGLAQACPEPCDCGEKYGFQIADCAYRDLESVPPGFPANVTTLSSLNRL
PGLPEGAFREVPLLQSLWLAHNEIRTVAAAGALASLSHLKSLDLSHNLI SDFAWSDLHNLSAL
QLLKMSNELTFIPRDAFRSLRALRSLQLNHNRLHTLAEGTFTPLTALSHLQINENPFDCTC
GIVWLKTWALTAVSIPEQDNIACTSPHVLKGTPLSRLPPLPCSAPSVQLSYQPSQDGAELR
PGFVLALHCDVDGQPAPQLHWHIQIPSGIVEITSPNVGTDGRALPGTPVASSQPRFQAFANG
SLLIPDFGKLEEGTYSCLATNELGSAESSVDVALATPGEGGEDTLGRRFHGKAVEGKGCYTV
DNEVQPSGPEDNVVVIYLSRAGNPEAAVAEGVPGQLPPGLLLLLGQSLLLFFFLTSF

Important features of the protein:

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 403-418

N-glycosylation sites.

amino acids 51-55, 120-124, 309-313

Tyrosine kinase phosphorylation site.

amino acids 319-326

N-myristoylation sites.

amino acids 14-20, 64-70, 92-98, 218-224, 294-300, 323-329,
334-340, 350-356, 394-400

Amidation site.

amino acids 355-359

Leucine rich repeats.

amino acids 51-74, 75-98, 99-122, 123-146, 147-170

Leucine rich repeat C-terminal domain.

amino acids 180-230

FIGURE 13

CCAGGCCGGGAGGCGACGCGCCAGCCGTCTAAACGGGAACAGCCCTGGCTGAGGGAGCTGC
AGCGCAGCAGAGTATCTGACGGCGCCAGGTTGCGTAGGTGCGGCACGAGGAGTTTCCCGGC
AGCGAGGAGGTCCTGAGCAGC**ATG**CCCCGGAGGAGCGCCTTCCCTGCCGCCGCGCTCTGGCT
CTGGAGCATCCTCCTGTGCCTGCTGGCACTGCGGGCGGAGGCCGGGCCGCCGAGGAGGAGA
GCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCATAGGATTTGAAGAAGATATC
CTGATTGTTTCAGAGGGGAAAATGGCACCTTTTACACATGATTTCAGAAAAGCGCAACAGAG
AATGCCAGCTATTCCTGTCAATATCCATTCCATGAATTTTACCTGGCAAGCTGCAGGGCAGG
CAGAATACTTCTATGAATTCCTGTCTTGGCTCCCTGGATAAAGGCATCATGGCAGATCCA
ACCGTCAATGTCCCTCTGCTGGGAACAGTGCCTCACAAGGCATCAGTTGTTCAAGTTGGTTT
CCCATGTCTTGGAAAACAGGATGGGGTGGCAGCATTTGAAGTGGATGTGATTGTTATGAATT
CTGAAGGCAACACCATTCTCCAAACACCTCAAAATGCTATCTTCTTTAAACATGTCAACAA
GCTGAGTGCCAGGCGGGTGCCGAAATGGAGGCTTTTGTAATGAAAGACGCATCTGCGAGTG
TCCTGATGGGTTCACGGACCTCACTGTGAGAAAGCCCTTTGTACCCACGATGTATGAATG
GTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCCACCTGGATTCTATGGAGTGAAGTGT
GACAAAGCAAACCTGCTCAACCACCTGCTTTAATGGAGGGACCTGTTTCTACCCTGGAAAATG
TATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAGCAAATGCCCACAACCCTGTC
GAAATGGAGGTAAATGCATTGGTAAAGCAAATGTAAGTGTTCCAAAGGTTACCAGGGAGAC
CTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACATGGAACCTGCCATGAACCCAA
CAAATGCCAATGTCAAGAAGGTTGGCATGGAAGACACTGCAATAAAAGGTACGAAGCCAGCC
TCATACATGCCCTGAGGCCAGCAGGCGCCAGCTCAGGCAGCACACGCCTTCACTTAAAAAG
GCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGG**TGA**ACTCCGACATCTGAAAC
GTTTTAAGTTACACCAAGTTCATAGCCTTTGTAAACCTTTCATGTGTTGAATGTTCAAATAA
TGTTTCATTACACTTAAGAATACTGGCCTGAATTTTATTAGCTTCATTATAAATCACTGAGCT
GATATTTACTCTTCCTTTTAAAGTTTTCTAAGTACGTCTGTAGCATGATGGTATAGATTTTCT
TGTTTCAGTGCTTTGGGACAGATTTTATATTATGTCAATTGATCAGGTTAAAATTTTCAGTG
TGTAAGTTGGCAGATATTTTCAAATTAACAATGCATTTATGGTGTCTGGGGGCAGGGGAACAT
CAGAAAGGTTAAATTGGGCAAAAATGCGTAAGTCACAAGAATTTGGATGGTGCAGTTAATGT
TGAAGTTACAGCATTTTCAGATTTTATTGTGAGATATTTAGATGTTTGTACATTTTAAAAA
TTGCTCTTAATTTTAACTCTCAATACAATATATTTTGACCTTACCATTATTCCAGAGATT
CAGTATTAATAAAAAAAAAAATTACACTGTGGTAGTGGCATTTAAACAATATAATATATTCTA
AACACAATGAAATAGGGAATATAATGTATGAACTTTTTGCATTGGCTTGAAGCAATATAATA
TATTGTAAACAAAACACAGCTCTTACCTAATAAACATTTTATACTGTTTGTATGTATAAAAT
AAAGGTGCTGCTTTAGTTTTTTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 14

MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESLYLWIDAHQARVLIGFEEDILIVSEGK
MAPFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFSLRSLDKGIMADPTVNVPLL
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPPGGC
RNGGFCNERRICECPDGFHGHCEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST
TCFNGGTCTFYPGKCICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSKPVC
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAERRDP
PESNYIW

Signal sequence.

amino acids 1-28

N-glycosylation sites.

amino acids 88-92, 245-249

Tyrosine kinase phosphorylation site.

amino acids 370-378

N-myristoylation sites.

amino acids 184-190, 185-191, 189-195, 315-321

ATP/GTP-binding site motif A (P-loop).

amino acids 285-293

EGF-like domain cysteine pattern signature.

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

AGAACCTCAGAAATGTGAGTTATTTGGGAATGGCTGTTTGTAATGTCCTTACGTAAGCCAA
 GAGGAGGTCTTGACTTGGGGTCCCAGGGGTACCGCAGATCCCAGGGACTGGAGCAGCACTAG
 CAAGCTCTGGAGGATGAGCCAGGAGTCTGGAATTGAGGCTGAGCCAAAGACCCCAGGGCCGT
 CTCAGTCTCATAAAAGGGGATCAGGCAGGAGGAGTTTGGGAGAAACCTGAGAAGGGCCTGAT
 TTGCGAGCATC**ATG**ATGGGCCTCTCCTTGGCCTCTGCTGTGCTCCTGGCCTCCCTCCTGAGTC
 TCCACCTTGGAAC TGCCACACGTGGGAGTGACATATCCAAGACCTGCTGCTTCCAATACAGC
 CACAAGCCCCTTCCCTGGACCTGGGTGCGAAGCTATGAATTCACCAGTAACAGCTGCTCCCA
 GCGGGCTGTGATATTCACTACCAAAAGAGGCAAGAAAGTCTGTACCCATCCAAGGAAAAAAT
 GGGTGCAAAAATACATTTCTTTACTGAAAAC TCCGAAACAATTG**TGA**CTCAGCTGAATTTTC
 ATCCGAGGACGCTTGGAACCCGCTCTTGGCTCTGCAGCCCTCTGGGGAGCCTGCGGAATCTT
 TTCTGAAGGCTACATGGACCCGCTGGGGAGGAGAGGGTGTTTCCTCCCAGAGTTACTTTAAT
 AAAGGTTGTTCATAGAGTTGAAA
 AA

16/75

FIGURE 16

MMGLSLASAVLLASLLSLHLGTATRGSDISKTC CFQYSHKPLPWTWVRSYEFTSNSCSQRAV
IFTTKRGKKVCTHPRKKWVQKYISLLKTPKQL

Important features of the protein:

Signal peptide:

amino acids 1-23

N-myristoylation sites.

amino acids 3-9, 26-32

Amidation site.

amino acids 68-72

Small cytokines (intecrine/chemokine).

amino acids 23-88

17/75

FIGURE 17

GCGAGAACCTTTGCACGCGCACAAACTACGGGGACGATTTCTGATTGATTTTTTGGCGCTTTCGATCCACCCTCC
TCCCTTCTC**ATG**GGGACTTTGGGGACAAAGCGTCCCGACCGCCTCGAGCGCTCGAGCAGGGCGCTATCCAGGAGC
CAGGACAGCGTCGGGAACCAGACCATGGCTCCTGGACCCCAAGATCCTTAAGTTCGTCGTCTTCATCGTCGCGG
TTCTGCTGCCGGTCCGGGTTGACTCTGCCACCATCCCCGGCAGGACGAAGTTCAGCAGACAGTGGCCCCA
CAGCAACAGAGGCGCAGCCTCAAGGAGGAGGAGTGTCCAGCAGGATCTCATAGATCAGAATATACTGGAGCCTG
TAACCCGTGCACAGAGGGTGTGGATTACACCATTTGCTTCCAACAATTTGCCTTCTTGCCCTGCTATGTACAGTTT
GTAATCAGGTCAAACAAATAAAAGTTTCTGTACCACGACAGAGACACCGTGTGTAGTGTGAAAAAGGAAGC
TTCCAGGATAAAAACTCCCTGAGATGTGCCGGACGTGTAGAACAGGGTGTCCAGAGGGATGGTCAAGGTCAG
TAATTGTACGCCCCGGAGTGACATCAAGTGCAAAATGAATCAGCTGCCAGTTCCTACTGGGAAAACCCAGCAG
CGGAGGAGACAGTGACCACCATCCTGGGGATGCTTGCCTCTCCCTATCACTACCTTATCATCATAGTGGTTTTA
GTCATCATTTTAGCTGTGGTTGTGGTTGGCTTTTCATGTGCGAAGAAATTCATTTCTTACCTCAAAGGCATCTG
CTCAGGTGGTGGAGGAGGTCCCGAACGTGTGCACAGAGTCCCTTTCCGGCGGCGTTCATGTCTTTCACGAGTTC
CTGGGGCGGAGGACAATGCCCGCAACGAGACCCCTGAGTAACAGATACTTGCAGCCACCCAGGTCTCTGAGCAG
GAAATCCAAGGTGAGGAGCTGGCAGAGCTAACAGGTGTGACTGTAGAGTCGCCAGAGGAGCCACAGCGTCTGCT
GGAACAGGCAGAAGCTGAAGGGTGTGAGAGGAGGAGGCTGCTGGTTCCAGTGAATGACGCTGACTCCGCTGACA
TCAGACCTTGCTGGATGCCTCGGCAACACTGGAAGAAGGACATGCAAGGAAACAATTCAGGACCAACTGGTG
GGCTCCGAAAAGCTCTTTTATGAAGAAGATGAGGCAGGCTCTGCTACGTCCTGCCTG**TGA**AAGAATCTCTTCAG
GAAACCAGAGCTTCCCTCATTTACCTTTTCTCTACAAAGGGAAGCAGCCTGGAAGAAACAGTCCAGTACTTGA
CCCATGCCCCAACAACTCTACTATCCAATATGGGGCAGCTTACCAATGGTCCTAGAACTTTGTAAACGCACTT
GGAGTAATTTTTATGAAATACTGCGTGTGATAAGCAAACGGGAGAAATTTATATCAGATTCTTGGCTGCATAGT
TATACGATTGTGTATTAAGGGTCGTTTTAGGCCACATGCGGTGGCTCATGCCTGTAATCCCAGCACTTTGATAG
GCTGAGGCAGGTGGATTGCTTGAGCTCGGGAGTTTGAGACCAGCCTCATCAACACAGTGAACATCCATCTCAAT
TTAAAAAGAAAAAAGTGGTTTTAGGATGTCTTTTGCAGTTCCTTCATCATGAGACAAGTCTTTTTCTGCTG
TTCTTATATTGCAAGCTCCATCTCTACTGGTGTGTGCATTTAATGACATCTAACTACAGATGCCGCACAGCCAC
AATGCTTTGCCTTATAGTTTTTTAACTTTAGAACGGGATTATCTTGTTATTACCTGTATTTTTCAGTTTCGGATA
TTTTTGACTTAATGATGAGATTATCAAGACGTAGCCCTATGCTAAGTCATGAGCATATGGACTTACGAGGGTTC
GACTTAGAGTTTTGAGCTTTAAGATAGGATTATGGGGCTTACCCCACTTAATTAGAGAAACATTTATATTG
CTTACTACTGTAGGCTGTACATCTCTTTCCGATTTTGTATAATGATGTAACATGGAAGAACTTTAGGAAAT
GCACTTATTAGGCTGTTTACATGGGTTGCCTGGATACAAATCAGCAGTCAAAAATGACTAAAAATATAACTAGT
GACGGAGGGAGAAATCCTCCCTCTGTGGGAGGCATTAAGTGCATTCCAGTTCCTCCCTCCTGCGCCCTGAGACTG
GACCAGGGTTTGATGGCTGGCAGCTTCTCAAGGGGCAGCTTGCTTACTTGTAAATTTTAGAGGTATATAGCCA
TATTTATTTATAAATAAATATTTATTTATTTATTTATAAGTAGATGTTTACATATGCCAGGATTTTGAAGAGC
CTGGTATCTTTGGGAAGCCATGTGTCTGGTTTGTGCTGGGACAGTCATGGGACTGCATCTTCCGACTTGTC
CACAGCAGATGAGGACAGTGAGAATTAAGTTAGATCCGAGACTGCCAAGAGCTTCTCTTTCAAGCGCCATTACA
GTTGAACGTTAGTGAATCTTGAGCCTCATTTGGGCTCAGGGCAGAGCAGGTGTTTATCTGCCCCGGCATCTGCC
ATGGCATCAAGAGGGAAGAGTGGACGGTGTCTTGGGAATGGTGTGAAATGGTTGCCGACTCAGGCATGGATGGGC
CCCTCTCGCTTCTGGTGGTCTGTGAAGTGAAGTCCCTGGGATGCCTTTTAGGGCAGAGATTCTGAGCTGCGTTT
TAGGGTACAGATTCCCTGTTTGGAGAGCTTGGCCCCCTCTGTAAGCATCTGACTCATCTCAGAGATATCAATTCT
TAAACACTGTGACAACGGGATCTAAAATGGCTGACACATTTGTCTTGTGTACGTTCCATTATTTTATTTAAA
AACCTCAGTAATCGTTTTAGCTTCTTTCCAGCAAACTCTTCTCCACAGTAGCCAGTCGTGGTAGGATAAATTA
CGGATATAGTCATTCTAGGGGTTTCAGTCTTTTCCATCTCAAGGCATTGTGTGTTTTGTTCGGGACTGGTTTG
GCTGGGACAAAGTTAGAAGTGCCTGAAGTTCGCACATTCAGATTGTTGTGTCCATGGAGTTTTAGGAGGGGATG
GCCTTTCCGCTTTCGCACTTCCATCCTCTCCCACTTCCATCTGGCGTCCACACCTTGTCCCTGCACTTCTG
GATGACACAGGGTGTGCTGCCTCCTAGTCTTTGCTTTGCTGGGCCCTTCTGTGCAGGAGACTTGGTCTCAAAG
CTCAGAGAGAGCCAGTCCGGTCCAGCTCCTTTGTCCCTTCTCAGAGGCCTTCTTGAAGATGCATCTAGACT
ACCAGCCTTATCAGTGTTTAAGCTTATTCCTTTAACATAAGCTTCCTGACAACATGAAATTTGTTGGGGTTTTTT
GGCGTTGGTTGATTTGTTTAGGTTTTGCTTTATACCCGGGCCAAATAGCACATAACACCTGGTTATATATGAAA
TACTCATATGTTTATGACCAAAATAAATATGAAACCTCATRTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 18

MGLWGQSVPTASSARAGRYPGARTASGTRPWLLDPKILKFVVFIVAVLLPVRVDSATIPRQD
EVPQQTVAPQQQRRSLKEEECPAGSHRSEYTGACNPCTEGVDYTIASNNLPSCLLCTVCKSG
QTNKSSCTTTTRDTVQCCEKGSFQDKNSPEMCRTCRTGCPRGMVKVSNCTPRSDIKCKNESAA
SSTGKTPAAEETVTTILGMLASPHYLI IIIVVLVIILAVVVVGFSCRKKFISYLGICSGGG
GGPERVHRVLFRRRSCPSRVPGAEDNARNETLSNRYLQPTQVSEQEIQGQELAEALTGVTVES
PEEPQRLLEQAEAEGCQRRRLLPVNDADSADISTLLDASATLEEGHAKETIQDQLVGSEKL
FYEEDEAGSATSCL

Important features of the protein:

Transmembrane domains:

amino acids 35-52, 208-230

N-glycosylation sites.

amino acids 127-131, 182-186, 277-281

Glycosaminoglycan attachment site.

amino acids 245-249

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 260-264

N-myristoylation sites.

amino acids 21-27, 86-92, 102-108, 161-167, 242-248, 270-276,
297-303, 380-386

ATP/GTP-binding site motif A (P-loop).

amino acids 185-193

TNFR/NGFR cysteine-rich region.

amino acids 99-139

FIGURE 19

GCGGCACCTGGAAG**ATG**CGCCCATTTGGCTGGTGGCCTGCTCAAGGTGGTGTTCGTGGTCTTC
GCCTCCTTGTGTGCCTGGTATTCGGGGTACCTGCTCGCAGAGCTCATTCCAGATGCACCCCT
GTCCAGTGCTGCCTATAGCATCCGCAGCATCGGGGAGAGGCCTGTCCTCAAAGCTCCAGTCC
CCAAAAGGCAAAAATGTGACCACTGGACTCCCTGCCCATCTGACACCTATGCCTACAGGTTA
CTCAGCGGAGGTGGCAGAAGCAAGTACGCCAAAATCTGCTTTGAGGATAACCTACTTATGGG
AGAACAGCTGGGAAATGTTGCCAGAGGAATAAACATTGCCATTGTCAACTATGTAAGTGGGA
ATGTGACAGCAACACGATGTTTTGATATGTATGAAGGCGATAACTCTGGACCGATGACAAAG
TTTATTCAGAGTGCTGCTCCAAAATCCCTGCTCTTCATGGTGACCTATGACGACGGAAGCAC
AAGACTGAATAACGATGCCAAGAATGCCATAGAAGCACTTGGAAGTAAAGAAATCAGGAACA
TGAAATTCAGGTCTAGCTGGGTATTTATTGCAGCAAAAGGCTTGGAACCTCCCTTCCGAAATT
CAGAGAGAAAAGATCAACCACTCTGATGCTAAGAACAACAGATATTCTGGCTGGCCTGCAGA
GATCCAGATAGAAGGCTGCATACCCAAAGAACGAAGCT**TGA**CACTGCAGGGTCTTGAGTAAAT
GTGTTCTGTATAAAACAAATGCAGCTGGAATCGCTCAAGAATCTTATTTTTCTAAATCCAACA
GCCCATATTTGATGAGTATTTTGGGTTTGTGTAAACCAATGAACATTTGCTAGTTGTATCA
AATCTTGGTACGCAGTATTTTATACCAGTATTTTATGTAGTGAAGATGTCAATTAGCAGGA
AACTAAAATGAATGGAAATTCTTAAAAAAAAA

FIGURE 20

MRPLAGGLLKVVVFVVFASLCAWYSGYLLAELIPDAPLSSAAYSIRSIGERPVLKAPVPKRQK
CDHWTPCPSDTYAYRLLSGGGRSKYAKICFEDNLLMGEQLGNVARGINIAIVNYVTGNVTAT
RCFDMYEGDNSGPMTKFIQSAAPKSLLFMVTYDDGSTRLNNDAKNAIEALGSKEIRNMKFRS
SWVFIAAKGLELPSEIQREKINHSDAKNNRYSGWPAEIQIEGCIPKERS

Important features of the protein:**Signal peptide:**

amino acids 1-20

N-glycosylation sites.

amino acids 120-124, 208-212

Glycosaminoglycan attachment site.

amino acids 80-84

N-myristoylation sites.

amino acids 81-87, 108-114, 119-125

FIGURE 21

CCGGGGAGGGGAGGGCCCGTCCCGCCCCCTCCCCGTCTCTCCCCGCCCCCTCCCCGTCCCTCCC
GCCGAAGCTCCGTCCCGCCCCGCGGGCCGGCTCCGCCCTCACCTCCCGGCCGCGGCTGCCCTC
TGCCCGGGTTGTCCAAGATGGAGGGCGCTCCACCGGGGTCGCTCGCCCTCCGGCTCCTGCTG
TTCGTGGCGCTACCCGCCTCCGGCTGGCTGACGACGGGCGCCCCCGAGCCGCCGCCGCTGTC
CGGAGCCCCACAGGACGGCATCAGAATTAATGTAACACTACACTGAAAGATGATGGGGACATAT
CTAAACAGCAGGTTGTTCTTAACATAACCTATGAGAGTGGACAGGTGTATGTAAATGACTTA
CCTGTAAATAGTGGTGTAAACCGAATAAGCTGTCAGACTTTGATAGTGAAGAATGAAAATCT
TGAAAATTTGGAGGAAAAAGAATATTTTGGAATTGTCAGTGTAAAGGATTTTAGTTCATGAGT
GGCCTATGACATCTGGTTCCAGTTTGCAACTAATTGTCATTCAAGAAGAGGTAGTAGAGATT
GATGGAAAACAAGTTCAGCAAAAGGATGTCAGTGAATTTGATATTTTAGTTAAGAACCGGGG
AGTACTCAGACATTCAAACCTATACCCTCCCTTTGGAAGAAAGCATGCTCTACTCTATTTCTC
GAGACAGTGACATTTTATTTACCCTTCCTAACCTCTCCAAAAAAGAAAGTGTTAGTTCACTG
CAAACCACTAGCCAGTATCTTATCAGGAATGTGGAAACCACTGTAGATGAAGATGTTTTACC
TGGCAAGTTACCTGAAACTCCTCTCAGAGCAGAGCCGCCATCTTCATATAAGGTAATGTGTC
AGTGGATGGAAAAGTTTAGAAAAGATCTGTGTAGGTTCTGGAGCAACGTTTTCCAGTATTC
TTTCAGTTTTTTGAACATCATGGTGGTTGGAATTACAGGAGCAGCTGTGGTAATAACCATCTT
AAAGGTGTTTTTCCCAGTTTCTGAATACAAAGGAATTCTTCAGTTGGATAAAGTGGACGTCA
TACCTGTGACAGCTATCAACTTATATCCAGATGGTCCAGAGAAAAGAGCTGAAAACCTTGAA
GATAAACATGTATTTTAAAACGCCATCTCATATCATGGACTCCGAAGTAGCCTGTTGCCTCC
AAATTTGCCACTTGAATATAATTTTCTTTAAATCGTT

amino acids 46-49, 64-67, 166-169, 191-194

FIGURE 23

CGTCTCTGCGTTTCGCC**ATG**CGTCCCGGGGCGCCAGGGCCACTCTGGCCTCTGCCCTGGGGGG
 CCCTGGCTTGGGCCGTGGGCTTCGTGAGCTCCATGGGCTCGGGGAACCCCGCGCCCGGTGGT
 GTTTGCTGGCTCCAGCAGGGCCAGGAGGCCACCTGCAGCCTGGTGCTCCAGACTGATGTAC
 CCGGGCCGAGTGCTGTGCCTCCGGCAACATTGACACCGCCTGGTCCAACCTCACCCACCCGG
 GGAACAAGATCAACCTCCTCGGCTTCTTGGGCCTTGTCCACTGCCTTCCCTGCAAAGATTTCG
 TGCAGCGGCGTGGAGTGCGGCCCGGGCAAGGCGTGCCGCATGCTGGGGGGCCGCCCGCGCTG
 CGAGTGCGCGCCCGACTGCTCGGGGCTCCCGGCGCGGCTGCAGGTCTGCGGCTCAGACGGCG
 CCACCTACCGCGACGAGTGCGGAGCTGCGCGCCGCGCTGCCGCGGCCACCCGGACCTGAGC
 GTCATGTACCGGGGCGCTGCCGCAAGTCTGTGAGCACGTGGTGTCGCCCGCGGCCACAGTC
 GTGCGTTCGTGGACCAGACGGGCAGCGCCACTGCGTGGTGTCGAGCGGGCGCCCTGCCCTG
 TGCCCTCCAGCCCCGGCCAGGAGCTTTGCGGCAACAACAACGTCACCTACATCTCCTCGTGC
 CACATGCGCCAGGCCACCTGCTTCCCTGGGCCGCTCCATCGGCGTGCGCCACGCGGGCAGCTG
 CGCAGGCACCCCTGAGGAGCCGCCAGGTGGTGAGTCTGCAGAAGAGGAAGAGAAGTTCGT**G**
GAGCCTGCAGGACAGGCCTGGGCCTGGTGCCCGAGGCCCCCATCATCCCTGTTATTTATT
 GCCACAGCAGAGTCTAATTTATATGCCACGGACACTCCTTAGAGCCCGGATTCGGACCACTT
 GGGGATCCCAGAACCTCCCTGACGATATCCTGGAAGGACTGAGGAAGGGAGGCCTGGGGGCC
 GGCTGGTGGGTGGGATAGACCTGCGTTCGGACACTGAGCGCCTGATTTAGGGCCCTTCTCT
 AGGATGCCCCAGCCCCTACCCTAAGACCTATTGCCGGGGAGGATTCCACACTTCCGCTCCTT
 TGGGGATAAACCTATTAATTATTGCTACTATCAAGAGGGCTGGGCATTCTCTGCTGGTAATT
 CCTGAAGAGGCATGACTGCTTTTCTCAGCCCCAAGCCTCTAGTCTGGGTGTGTACGGAGGGT
 CTAGCCTGGGTGTGTACGGAGGGTCTAGCCTGGGTGAGTACGGAGGGTCTAGCCTGGGTGAG
 TACGGAGGGTCTAGCCTGGGTGAGTACGGAGGGTCTAGCCTGGGTGTGTATGGAGGATCTAG
 CCTGGGTGAGTATGGAGGGTCTAGCCTGGGTGAGTATGGAGGGTCTAGCCTGGGTGTGTATG
 GAGGGTCTAGCCTGGGTGAGTATGGAGGGTCTAGCCTGGGTGTGTATGGAGGGTCTAGCCTG
 GGTGAGTATGGAGGGTCTAGCCTGGGTGTGTACGGAGGGTCTAGTCTGAGTGCGTGTGGGGA
 CCTCAGAACACTGTGACCTTAGCCCAGCAAGCCAGGCCCTTCATGAAGGCCAAGAAGGCTGC
 CACCATTTCCCTGCCAGCCCAAGAACTCCAGCTTCCCCACTGCCTCTGTGTGCCCTTTGCGT
 CCTGTGAAGGCCATTGAGAAATGCCAGTGTGCCCCCTGGGAAAGGGCACGGCCTGTGCTCC
 TGACACGGGCTGTGCTTGGCCACAGAACCACCCAGCGTCTCCCCTGCTGCTGTCCACGTCAG
 TTCATGAGGCAACGTCGCGTGGTCTCAGACGTGGAGCAGCCAGCGGCAGCTCAGAGCAGGGC
 ACTGTGTCCGGCGGAGCCAAGTCCACTCTGGGGGAGCTCTGGCGGGGACCACGGGCCACTGC
 TCACCCACTGGCCCCGAGGGGGGTGTAGACGCCAAGACTCACGCATGTGTGACATCCGGAGT
 CCTGGAGCCGGGTGTCCAGTGGCACCCTAGGTGCCTGCTGCCTCCACAGTGGGGTTCACA
 CCCAGGGCTCCTTGGTCCCCCACAACTGCCCGGGCCAGGCCTGCAGACCCAGACTCCAGCC
 AGACCTGCCTCACCCACCAATGCAGCCGGGGCTGGCGACACCAGCCAGGTGCTGGTCTTGGG
 CCAGTTCTCCACGACGGCTCACCCCTCCCTCCATCTGCGTTGATGCTCAGAATCGCCTACC
 TGTGCCTGCGTGTAAACCACAGCCTCAGACCAGCTATGGGGAGAGGACAACACGGAGGATAT
 CCAGCTTCCCCGGTCTGGGGTGAGGAATGTGGGGAGCTTGGGCATCCTCCTCCAGCCTCCTC
 CAGCCCCCAGGCAGTGCCCTTACCTGTGGTGCCAGAAAAGTGCCCCCTAGGTTGGTGGGTCTA
 CAGGAGCCTCAGCCAGGCAGCCACCCACCCCTGGGGCCCTGCCTCACCAAGGAAATAAAGA
 CTAAGCCATAAAAAAA

24/75

FIGURE 24

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62306

<subunit 1 of 1, 263 aa, 1 stop

<MW: 27663, pI: 6.77, NX(S/T): 2

MRPGAPGPLWPLPWGALAWAVGVSSMSGGNPAPGGVCWLQQGQEATCSLVLQTDVTRAEC
ASGNIDTAWSNLTHPGNKINLLGFLGLVHCLPCKDSCDGVCEGPGKACRMLGGRPRCECAPD
CSGLPARLQVCGSDGATYRDECELRAARCRGHPDLSVMYRGRCRKSCHEHVVCPRPQSCVVDQ
TGSAHCVVCRAAPCPVPSSPGQELCGNNNVITYISSCHMRQATCFLGRSIGVRHAGSCAGTPE
EPPGGESAEEEEENFV

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation sites.

amino acids 73-77, 215-219

Osteonectin domain proteins.

amino acids 97-130, 169-202

FIGURE 25

TGCAGAGCTTGTGGAGGCC**ATG**GGGCGCGTCGTCGCGGAGCTCGTCTCCTCGCTGCTGGGGT
TGTGGCTGTTGCTGTGCAGCTGCGGATGCCCCGAGGGCGCCGAGCTGCGTGCTCCGCCAGAT
AAAATCGCGATTATTGGAGCCGGAATTGGTGGCACTTCAGCAGCCTATTACCTGCGGCAGAA
ATTTGGGAAAGATGTGAAGATAGACCTGTTTGAAGAGAAGAGGTCGGGGGGCCGCTGGCTA
CCATGATGGTGCAGGGGCAAGAATACGAGGCAGGAGGTTCTGTTCATCCATCCTTTAAATCTG
CACATGAAACGTTTTGTCAAAGACCTGGGTCTCTCTGCTGTTTCAGGCCTCTGGTGGCCTACT
GGGGATATATAATGGAGAGACTCTGGTATTTGAGGAGAGCAACTGGTTCATAATTAACGTGA
TTAAATTAGTTTGGCGCTATGGATTTCAATCCCTCCGTATGCACATGTGGGTAGAGGACGTG
TTAGACAAGTTCATGAGGATCTACCGCTACCACTCTCATGACTATGCCTTCAGTAGTGTCGA
AAAATTACTTCATGCTCTAGGAGGAGATGACTTCCTTGGAATGCTTAATCGAACACTTCTTG
AAACCTTGCAAAGGCCGGCTTTTCTGAGAAGTTCCTCAATGAAATGATTGCTCCTGTTATG
AGGGTCAATTATGGCCAAAGCACGGACATCAATGCCTTTGTGGGGGCGGTGTCACTGTCCTG
TTCTGATTCTGGCCTTTGGGCAGTAGAAGGTGGCAATAAACTTGTTTGTCTAGGGCTTCTGC
AGGCATCCAAAAGCAATCTTATATCTGGCTCAGTAATGTACATCGAGGAGAAAACAAAGACC
AAGTACACAGGAAATCCAACAAAGATGTATGAAGTGGTCTACCAAATTGGAACGTGAGACTCG
TTCAGACTTCTATGACATCGTCTTGGTGGCCACTCCGTTGAATCGAAAAATGTCGAATATTA
CTTTTCTCAACTTTGATCCTCCAATTGAGGAATTCATCAATATTATCAACATATAGTGACA
ACTTTAGTTAAGGGGGAATTGAATACATCTATCTTTAGCTCTAGACCCATAGATAAAATTTGG
CCTTAATACAGTTTTTAACCACTGATAATTCAGATTTGTTTCATTAACAGTATTGGGATTGTGC
CCTCTGTGAGAGAAAAGGAAGATCCTGAGCCATCAACAGATGGAACATATGTTTGGAAAGATC
TTTTCCCAAGAACTCTTACTAAAGCACAAATTTTAAAGCTCTTTCTGTCCTATGATTATGC
TGTGAAGAAGCCATGGCTTGCATATCCTCACTATAAGCCCCCGGAGAAATGCCCTCTATCA
TTCTCCATGATCGACTTTATTACCTCAATGGCATAGAGTGTGCAGCAAGTGCCATGGAGATG
AGTGCCATTGCAGCCCACAACGCTGCACTCCTTGCCTATCACCGCTGGAACGGGCACACAGA
CATGATTGATCAGGATGGCTTATATGAGAACTTAAACTGAACTAT**TGA**AGTGACACACTCC
TTTTTCCCCTCCTAGTTCCAAATGACTATCAGTGGCAAAAAGAACAAAATCTGAGCAGAGA
TGATTTTGAACCAGATATTTTGCCATTATCATTGTTTAATAAAAGTAATCCCTGCTGGTCAT
AGGAAAAAAAAAAAAA

26/75

FIGURE 26

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62880

<subunit 1 of 1, 505 aa, 1 stop

<MW: 56640, pI: 6.10, NX(S/T): 4

MGRVVAELVSSLLGLWLLLCSCGCPEGAE LRAPPDKIAIIGAGIGGTS AAYYL RQKFGKDVK
IDLFEREEVGGRLATMMVQGQ EYEAGGSVIHPLNLHMKRFVKDLGLSAVQASGGLLGIYNGE
TLVFEE SNWFI INVIKLVWRYGFQSLRMH MWVEDVLDKFMRIYRYQSHDYAFSSVEKLLHAL
GGDDFLGMLNRTLLET LQKAGFSEKFLNEMIAPVMRVNYGQSTDINA FVGAVSLSCSDSGLW
AVEGGNKLVC SGLLQASKSNLISGSVMYIEEKT KTKYTGNPTKMYEVVYQIGTETRSDFYDI
VLVATPLNRKMSNITFLNFDPPIEEFHQYYQHIVTTLVKGELNTSIFSSRPIDKFG LNTVLT
TDNSDLFINSIGIVPSVREKEDPEPSTDGTYVWKIFSQETLT KAQILKLFLSYDYAVKKPWL
AYPHYKPPEKCP SIILHDRLYY LNGIECAASAMEMSAIAAHNAALLAYHRWNGHTDMIDQDG
LYEKLKTEL

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation sites.

amino acids 196-200, 323-327, 353-357

Tyrosine kinase phosphorylation site.

amino acids 291-298

N-myristoylation sites.

amino acids 23-29, 41-47, 43-49, 45-51, 46-52, 72-78, 115-121,
119-125, 260-266, 384-390, 459-465

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 12-23, 232-243

FIGURE 27

CATTTCCAACAAGAGCACTGGCCAAGTCAGCTTCTTCTGAGAGAGTCTCTAGAAGAC**ATG**AT
GCTACACTCAGCTTTGGGTCTCTGCCTCTTACTCGTCACAGTTTCTTCCAACCTTGCCATTG
CAATAAAAAAGGAAAAGAGGCCTCCTCAGACACTCTCAAGAGGATGGGGAGATGACATCACT
TGGGTACAACTTATGAAGAAGGTCTCTTTTATGCTCAAAAAAGTAAGAAGCCATTAATGGT
TATTCATCACCTGGAGGATTGTCAATACTCTCAAGCACTAAAGAAAGTATTTGCCCAAATG
AAGAAATACAAGAAATGGCTCAGAATAAGTTCATCATGCTAAACCTTATGCATGAAACCACT
GATAAGAATTTATCACCTGATGGGCAATATGTGCCTAGAATCATGTTTGTAGACCCTTCTTT
AACAGTTAGAGCTGACATAGCTGGAAGATACTCTAACAGATTGTACACATATGAGCCTCGGG
ATTTACCCCTATTGATAGAAAACATGAAGAAAGCATTAAGACTTATTCAGTCAGAGCTA**TAA**
GAGATGATGGAAAAAAGCCTTCACTTCAAAGAAGTCAAATTTTCATGAAGAAAACCTCTGGCA
CATTGACAAATACTAAATGTGCAAGTATATAGATTTTGTAATATTACTATTTAGTTTTTTTA
ATGTGTTTGCAATAGTCTTATTAAAATAAATGTTTTTTTAAATCTGA

28/75

FIGURE 28

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64896

<subunit 1 of 1, 166 aa, 1 stop

<MW: 19171, pI: 8.26, NX(S/T): 1

MMLHSALGLCLLLVTVSSNLAIKKEKRPPQTLSRGWGDDITWVQTYEEGLFYAQSKKPL
MVIHHLEDCQYSQALKKVFQAQNEEIQEMAQNKFIMLNLMHETTDKNLSPDGQYVPRIMFVDP
SLTVRADIAGRYSNRLYTYEPRDLPLLIENMKKALRLIQSEL

Important features:

Signal peptide:

amino acids 1-23

N-myristoylation site.

amino acids 51-57

29/75

FIGURE 29

TAAACAGCTACAATATTCCAGGGCCAGTCACTTGCCATTTCTCATAACAGCGTCAGAGAGA
AAGAACTGACTGAAACGTTTGAG**ATGA**AAGAAAGTTCTCCTCCTGATCACAGCCATCCTTGGCA
GTGGCTGTTGGTTTCCCAGTCTCTCAAGACCAGGAACGAGAAAAAAGAAGTATCAGTGACAG
CGATGAATTAGCTTCAGGGTTTTTTGTGTTCCCTTACCCATATCCATTTGCGCCCACTTCCAC
CAATTCCATTTCCAAGATTTCCATGGTTTAGACGTAATTTTCTTATTCCAATACCTGAATCT
GCCCCTACAACCTCCCCTTCCTAGCGAAAAG**TAA**ACAAGAAGGATAAGTCACGATAAACCTGG
TCACCTGAAATTGAAATTGAGCCACTTCCTTGAAGAATCAAAATTCCTGTTAATAAAAGAAA
AACAAATGTAATTGAAATAGCACACAGCATTCTCTAGTCAATATCTTTAGTGATCTTCTTTA
ATAAACATGAAAGCAAAGATTTTGGTTTCTTAATTTCCACA

30/75

FIGURE 30

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71290
><subunit 1 of 1, 85 aa, 1 stop
><MW: 9700, pI: 9.55, NX(S/T): 0
MKKVLLLITAILAVAVGFPVSQDQEREKRSISDSDELASGFFVFPYPYPFRPLPPIPFFPRFP
WFRRNFPIPIPIESAPTTPLPSEK
```

Important features of the protein:

Signal peptide:

amino acids 1-17

Homologous region to B3-hordein:

amino acids 47-85

FIGURE 31

CGGACGCGTGGGCGGGCGCGCCGGGAGGGACCGGCGGCGGC**ATG**GGGCCGGGGGCCCTGGGAT
 GCGGGCCCCGTCTCGCCGCTGCTGCCGCTGTTGCTGCTGCTCGGCCTGGCCCCGGGCGCCGC
 GGGAGCGCCGGGCCCCGACGGTTTAGACGTCTGTGCCACTTGCCATGAACATGCCACATGCC
 AGCAAAGAGAAGGGAAGAAGATCTGTATTTGCAACTATGGATTTGTAGGGAACGGGAGGACT
 CAGTGTGTTGATAAAAATGAGTGCCAGTTTGGAGCCACTCTTGTCTGTGGGAACACACATC
 TTGCCACAACACCCCCGGGGGCTTCTATTGCATTTGCCTGGAAGGATATCGAGCCACAAACA
 ACAACAAGACATTTCATTCCCAACGATGGCACCTTTTGTACAGACATAGATGAGTGTGAAGTT
 TCTGGCCTGTGCAGGCATGGAGGGCGATGCGTGAACACTCATGGGAGCTTTGAATGCTACTG
 TATGGATGGATACTTGCCAAGGAATGGACCTGAACCTTTCCACCCGACCACCGATGCCACAT
 CATGCACAGAAATAGACTGTGGTACCCCTCCTGAGGTTCCAGATGGCTATATCATAGGAAAT
 TATACGTCTAGTCTGGGCAGCCAGGTTCTGTTATGCTTGCAGAGAAGGATTCTTCAGTGTTC
 AGAAGATACAGTTTCAAGCTGCACAGGCCTGGGCACATGGGAGTCCCCAAAATTACATTGCC
 AAGAGATCAACTGTGGCAACCCCTCCAGAAATGCGGCACGCCATCTTGGTAGGAAATCACAGC
 TCCAGGCTGGGCGGTGTGGCTCGCTATGTCTGTCAAGAGGGCTTTGAGAGCCCTGGAGGAAA
 GATCACTTCTGTTTGCACAGAGAAAGGCACCTGGAGAGAAAGTACTTTAACATGCACAGAAA
 TTCTGACAAAGATTAATGATGTATCACTGTTTAAATGATACCTGTGTGAGATGGCAAATAAAC
 TCAAGAAGAATAAACCCCAAGATCTCATATGTGATATCCATAAAAGGACAACGGTTGGACCC
 TATGGAATCAGTTTCGTGAGGAGACAGTCAACTTGACCACAGACAGCAGGACCCCAAGTGT
 GCCTAGCCCTGTACCCAGGCACCAACTACACCGTGAACATCTCCACAGCACCTCCCAGGCGC
 TCGATGCCAGCCGTCATCGGTTTCCAGACAGCTGAAGTTGATCTCTTAGAAGATGATGGAAG
 TTTCAATATTTCAATATTTAATGAAACTTGTGTTGAAATTGAACAGGCGTTCTAGGAAAGTTG
 GATCAGAACACATGTACCAATTTACCGTTCTGGGTCAGAGGTGGTATCTGGCTAACTTTTCT
 CATGCAACATCGTTTAACTTCACAACGAGGGAACAAGTGCCTGTAGTGTGTTTGGATCTGTA
 CCCTACGACTGATTATACGGTGAATGTGACCCTGCTGAGATCTCCTAAGCGGCACTCAGTGC
 AAATAACAATAGCAACTCCCCCAGCAGTAAACAGACCATCAGTAACATTTTCAGGATTTAAT
 GAAACCTGCTTGAGATGGAGAAGCATCAAGACAGCTGATATGGAGGAGATGTATTTATTCCA
 CATTTGGGGCCAGAGATGGTATCAGAAGGAATTTGCCCAGGAAATGACCTTTAATATCAGTA
 GCAGCAGCCGAGATCCCGAGGTGTGCTTGGACCTACGTCCGGGTACCAACTACAATGTCAGT
 CTCCGGGCTCTGTCTTCGGAACCTTCTGTGGTTCATCTCCCTGACAACCCAGATAACAGAGCC
 TCCCTTCCCGGAAGTAGAATTTTTTTACGGTGCACAGAGGACCTCTACCACGCCTCAGACTGA
 GGAAAGCCAAGGAGAAAAATGGACCAATCAGTTCATATCAGGTGTTAGTGCTTCCCTTGGCC
 CTCCAAAGCACATTTTCTTGTGATTCTGAAGGCGCTTCCCTCCTTCTTTAGCAACGCCTCTGA
 TGCTGATGGATACGTGGCTGCAGAACTACTGGCCAAAGATGTTCCAGATGATGCCATGGAGA
 TACCTATAGGAGACAGGCTGTACTATGGGGAATATTATAATGCACCCCTTGAAAAGAGGGAGT
 GATTACTGCATTATATTACGAATCACAAGTGAATGGAATAAGGTGAGAAGACACTCCTGTGC
 AGTTTGGGCTCAGGTGAAAGATTCGTCACTCATGCTGCTGCAGATGGCGGGTGTGGACTGG
 GTTCCCTGGCTGTTGTGATCATTCTCACATTCCTCTCCTTCTCAGCGGTG**TGAT**GGCAGATG
 GACACTGAGTGGGGAGGATGCACTGCTGCTGGGCAGGTGTTCTGGCAGCTTCTCAGGTGCC
 GCACAGAGGCTCCGTGTGACTTCCGTCCAGGGAGCATGTGGGCCTGCAACTTTCTCCATTCC
 CAGCTGGGCCCCATTCTGGATTTAAGATGGTGGCTATCCCTGAGGAGTCACCATAAGGAGA
 AACTCAGGAATTCTGAGTCTTCCCTGCTACAGGACCAGTTCTGTGCAATGAACCTTGAGACT
 CCTGATGTACACTGTGATATTGACCGAAGGCTACATACAGATCTGTGAATCTTGGCTGGGAC
 TTCTCTGAGTGTGCTGAGGGTCACTCCTCTAGACATTGACTGCAAGAGAATCTCTGCA
 ACCTCCTATATAAAAGCATTTCTGTAAATTCATTGAGAATCCATTCTTTACAATATGCAGTG
 AGATGGGCTTAAGTTTGGGCTAGAGTTTGACTTTATGAAGGAGGTCAATTGAAAAAGAGAAC
 GTGACGTAGGCAAATGTTTCAAGCACTTTAGAAACAGTACTTTTCTATAATTAGTTGATAT
 ACTAATGAGAAAATATACTAGCCTGGCCATGCCAATAAGTTTCTGCTGTGTCTGTTAGGCA
 GCATTGCTTTGATGCAATTTCTATTGTCCTATATATTCAAAGTAATGTCTACATTCCAGTA
 AAAATATCCCGTAATTAAAAA

FIGURE 32

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96031
><subunit 1 of 1, 747 aa, 1 stop
><MW: 82710, pI: 6.36, NX(S/T): 18
MGRGPWDAGPSRRLPLLLLLLGLARGAAGAPGPDGLDVCATCHEHATCQQQREGKKICICNYG
FVGNGRTQCVDKNECQFGATLVCGNHTSCHNTPGGFYCICLEGYRATNNNKTFIPNDGTFCT
DIDECEVSGLCRHGGRCVNTHGSEFCYCMDGYLPRNGPEPFHPTTDATSCTEIDCGTPPEVP
DGYIIGNYTSSLGSQVRYACREGFFSVPEDTVSSCTGLGTWESPKLHCQEINCGNPPEMRHA
ILVGNHSSRLGGVARYVCQEGFESPGGKITSVCTEKGTWRESTLTCTEILTKINDVSLFNDT
CVRWQINSRRINPKISYVISIKGQRLDPMESVREETVNLTDSRTPEVCLALYPGTNYTVNI
STAPPRRSMFAVIGFQTAEVDLLEDDGSFNISIFNETCLKLNRRSRKVGSEHMYQFTVLGQR
WYLANFSHATSFNFTTREQVPVVCCLDLYPTTDYTVNVTLLRSPKRHSVQITIATPPAVKQTI
SNISGFNETCLRWSIKTADMEEMYLFIHWGQRWYQKEFAQEMTFNIISSSSRDPEVCLDLRP
GTNYNVSLRALSSELPVVISLTQTITEPPLPEVEFFTVHRGPLPRLRLRKAKEKNGPISSYQ
VLVLPLALQSTFSCDSEGASSFFSNASDADGYVAAELLAKDVPDDAMEIPIGDRLYYGEYYN
APLKRGSDYCIILRITSEWNKVRRHSCAVWAQVKDSSLMLLQMAGVGLGSLAVVIILTFLSF
SAV
```

Important features of the protein:**Signal peptide:**

amino acids 1-29

Transmembrane domain:

amino acids 718-740

N-glycosylation sites.

amino acids 87-91, 112-116, 193-197, 253-257, 308-312, 348-352, 367-371, 371-375, 402-406, 407-411, 439-443, 447-451, 470-474, 498-502, 503-507, 542-546, 563-567, 645-649

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 478-482, 686-690, 705-709

Tyrosine kinase phosphorylation site.

amino acids 419-427

N-myristoylation sites.

amino acids 22-28, 35-41, 65-71, 86-92, 96-102, 120-126, 146-152, 192-198, 252-258, 274-280, 365-371, 559-565, 688-694, 727-733.

Amidation site.

amino acids 52-56

Aspartic acid and asparagine hydroxylation sites.

amino acids 91-103, 141-153.

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 624-635

Cytochrome c family heme-binding site signature.

amino acids 39-45

Calcium-binding EGF-like domain proteins pattern proteins.

amino acids 85-106, 135-156

Receptor tyrosine kinase class V proteins:

amino acids 389-422

FIGURE 33

GGAAAAGGTACCCGCGAGAGACAGCCAGCAGTTCTGTGGAGCAGCGGTGGCCGGCTAGG**ATG**
GGCTGTCTCTGGGGTCTGGCTCTGCCCCTTTTCTTCTTCTGCTGGGAGGTTGGGGTCTCTGG
GAGCTCTGCAGGCCCCAGCACCCGCGAGAGCAGACACTGCGATGACAACGGACGACACAGAAG
TGCCCGCTATGACTCTAGCACCCGGGCCACGCCGCTCTGGAAACTCAAACGCTGAGCGCTGAG
ACCTCTTCTAGGGCCTCAACCCAGCCGGCCCCATTCCAGAAGCAGAGACCAGGGGAGCCAA
GAGAATTTCCCCTGCAAGAGAGACCAGGAGTTTCACAAAAACATCTCCCAACTTCATGGTGC
TGATCGCCACCTCCGTGGAGACATCAGCCGCCAGTGGCAGCCCCGAGGGAGCTGGAATGACC
ACAGTTCAGACCATCACAGGCAGTGATCCCGAGGAAGCCATCTTTGACACCCTTTGCACCGA
TGACAGCTCTGAAGAGGCAAAGACACTCACAATGGACATATTGACATTGGCTCACACCTCCA
CAGAAGCTAAGGGCCTGTCCTCAGAGAGCAGTGCCCTCTCCGACGGCCCCCATCCAGTCATC
ACCCCGTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCC
GTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCCGTCAT
GGTCCCCGGGATCTGATGTCACTCTCCTCGCTGAAGCCCTGGTGACTGTCAAAACATCGAG
GTTATTAATTGCAGCATCACAGAAATAGAAACAACAACTTCCAGCATCCCTGGGGCCTCAGA
CATAGATCTCATCCCCACGGAAGGGGTGAAGGCCTCGTCCACCTCCGATCCACCAGCTCTGC
CTGACTCCACTGAAGCAAAACCACACATCACTGAGGTACAGCCTCTGCCGAGACCCTGTCC
ACAGCCGGCACCACAGAGTCAGCTGCACCTCATGCCACGGTTGGGACCCCACTCCCCACTAA
CAGCGCCACAGAAAGAGAAGTGACAGCACCCGGGGCCACGACCCTCAGTGGAGCTCTGGTCA
CAGTTAGCAGGAATCCCCTGGAAGAAACCTCAGCCCTCTCTGTTGAGACACCAAGTTACGTC
AAAGTCTCAGGAGCAGCTCCGGTCTCCATAGAGGCTGGGTGAGCAGTGGGCAAAACAACCTTC
CTTTGCTGGGAGCTCTGCTTCCTCCTACAGCCCCTCGGAAGCCGCCCTCAAGAACTTCACCC
CTTCAGAGACACCGACCATGGACATCGCAACCAAGGGGGCCCTTCCCCACCAGCAGGGACCCCT
CTTCCTTCTGTCCCTCCGACTACAACCAACAGCAGCCGAGGGACGAACAGCACCTTAGCCAA
GATCACAACCTCAGCGAAGACCACGATGAAGCCCCAACAGCCACGCCACGACTGCCCGGAC
GAGGCCGACCACAGACG**TGA**GTGCAGGTGAAAATGGAGGTTTCCTCCTCCTGCGGCTGAGTG
TGGCTTCCCCGGAAGACCTCACTGACCCCAAGAGTGGCAGAAAGGCTGATGCAGCAGCTCCAC
CGGGAACCTCCACGCCCACGCGCCTCACTTCCAGGTCTCCTTACTGCGTGTGAGGAGAGGCTA
ACGGACATCAGCTGCAGCCAGGCATGTCCCGTATGCCAAAAGAGGGTGCTGCCCTAGCCTG
GGCCCCACCGACAGACTGCAGCTGCGTTACTGTGCTGAGAGGTACCCAGAAGGTTCCCATG
AAGGGCAGCATGTCCAAGCCCCTAACCCAGATGTGGCAACAGGACCCCTCGCTCACATCCAC
CGGAGTGTATGTATGGGGAGGGGCTTCACCTGTTCCAGAGGTGTCCTTGGACTCACCTTGG
CACATGTTCTGTGTTTCAGTAAAGAGAGACCTGATACCCATCTGTGTGCTTCCATCCTGCA
TTAAAATTCACTCAGTGTGGCCCCAAAAAAA

FIGURE 34

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108722
><subunit 1 of 1, 482 aa, 1 stop
><MW: 49060, pI: 4.74, NX(S/T): 4
MGCLWGLALPLFFFCWEVGVSGSSAGPSTRRADTAMTTDDTEVPAMTLAPGHAALETQTLA
ETSSRASTPAGPIPEAETRGAKRISPARETRSFTKTSPNFMVLIATSVETSAASGSPEGAGM
TTVQTTITGSDPEEAIFDTLCTDDSSEEAKTLTMDILTTLAHTSTEAKGLSSESSASSDGPHPV
ITPSRASESSASSDGPHPVITPSRASESSASSDGPHPVITPSWSPGSDVTLLEALVTVTNI
EVINCSITEIETTTSSIPGASDIDLIPTEGVKASSTSDPPALPDSTEAKPHITEVTASAETL
STAGTTESAAPHATVGTPLPTNSATEREVTAPGATTLGALVTVSRNPLEETSALSVETPSY
VKVSGAAPVSI EAGSAVGKTTSFAGSSASSYSPSEAALKNFTPSETPTMDIATKGPFPTSRD
PLPSVPPTTTNSSRGTNSTLAKITTSAKTTMKPQQPRPRLPGRGRPQT
```

Important features of the protein:**Signal peptide:**

amino acids 1-25

N-glycosylation sites.

amino acids 252-256, 445-449, 451-455

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 84-88

N-myristoylation sites.

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284,
314-320, 349-355, 386-392, 397-403, 449-455

ATP/GTP-binding site motif A (P-loop).

amino acids 385-393

FIGURE 35

GCCTCTGAATTGTTGGGCAGTCTGGCAGTGGAGCTCTCCCCGGTCTGACAGCCACTCCAGAG
GCC**ATG**CCTTCGTTTCTTGCCAGATTTGGCTTTCAGCTTCCTGTAAATTCTGGCTTTGGGCCA
GGCAGTCCAATTTCAAGAATATGTCTTTCTCCAATTTCTGGGCTTAGATAAGGCGCCTTCAC
CCCAGAAGTTCCAACCTGTGCCTTATATCTTGAAGAAAATTTTCCAGGATCGCGAGGCAGCA
GCGACCACTGGGGTCTCCCGAGACTTATGCTACGTAAAGGAGCTGGGCGTCCGCGGGAATGT
ACTTCGCTTTCTCCCAGACCAAGGTTTCTTTCTTTACCCAAAGAAAATTTCCCAAGCTTCCT
CCTGCCTGCAGAAGCTCCTCTACTTTAACCTGTCTGCCATCAAAGAAAGGGAACAGTTGACA
TTGGCCCAGCTGGGCCTGGACTTGGGGCCCAATTCTTACTATAACCTGGGACCAGAGCTGGA
ACTGGCTCTGTTCTGGTTCAGGAGCCTCATGTGTGGGGCCAGACCACCCCTAAGCCAGGTA
AAATGTTTGTGTTGCGGTCAGTCCCATGGCCACAAGGTGCTGTTCACTTCAACCTGCTGGAT
GTAGCTAAGGATTGGAATGACAACCCCCGGAAAAATTTCTGGGTATTCTGAGATACTGGT
CAAAGAAGATAGAGACTCAGGGGTGAATTTTCAGCCTGAAGACACCTGTGCCAGACTAAGAT
GCTCCCTTCATGCTTCCCTGCTGGTGGTGACTCTCAACCCTGATCAGTGCCACCCTTCTCGG
AAAAGGAGAGCAGCCATCCCTGTCCCCAAGCTTTCTTGTAAGAACCTCTGCCACCGTCACCA
GCTATTCATTAACCTCCGGGACCTGGGTGGCACAAGTGATCATTGCCCCAAGGGGTTC
TGGCAAATTACTGCCATGGAGAGTGTCCCTTCTCACTGACCATCTCTCTCAACAGCTCCAAT
TATGCTTTCATGCAAGCCCTGATGCATGCCGTTGACCCAGAGATCCCCCAGGCTGTGTGTAT
CCCCACCAAGCTGTCTCCCATTTCCATGCTCTACCAGGACAATAATGACAATGTCATTCTAC
GACATTATGAAGACATGGTAGTCGATGAATGTGGGTGTGGG**TAG**GATGTCAGAAATGGGAAT
AGAAGGAGTGTTCTTAGGGTAAATCTTTTAATAAACTACCTATCTGGTTTATGACCACTTA
GATCGAAATGTC

FIGURE 36

MLRFLPDLAFSFLILALGQAVQFQEYVFLQFLGLDKAPSPQKFQVPVYILKKIFQDREAAA
TTGVSRDLCYVKELGVRGNVLRFLPDQGFFLYPKKISQASSCLQKLLYFNLSAIKEREQLTL
AQLGLDLGPNSYYNLGPELELALFLVQEPHVWGQTTPKPGKMFVLRVWPWPQGAVHFNLLDV
AKDWNDNPRKNFGLFLEILVKEDRDSGVNFQPEDTCARLRCSLHASLLVVTLNPDQCHPSRK
RRAAIPVPKLSCKNLCHRHQLFINFRDLGWHKWIIAPKGFMANYPCHGECPPSLTISLNSSNY
AFMQALMHAVDPEIPQAVCIPTKLSPISMLYQDNNDNVILRHYEDMVVDECGCG

Important features of the protein:**Signal peptide:**

amino acids 1-21

N-glycosylation sites.

amino acids 112-116, 306-310

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 96-100

N-myristoylation site.

amino acids 77-83

TGF-beta family proteins.

amino acids 264-299, 327-341, 345-364

FIGURE 37

CACTTTCTCCCTCTCTTCTTTACTTTTCGAGAAACCGCGCTTCCGCTTCTGGTCGCAGAGAC
CTCGGAGACCGCGCCGGGGAGACGGAGGTGCTGTGGGTGGGGGGGACCTGTGGCTGCTCGTA
CCGCCCCCACCCTCCTCTTCTGCACTGCCGTCTCCGGAAGACCTTTTCCCTGCTCTGTT
TCCTTCACCGAGTCTGTGCATCGCCCCGACCTGGCCGGGAGGAGGCTTGGCCGGCGGGAGA
TGCTCTAGGGGCGGCGGGGAGGAGCGGCCGGCGGGACGGAGGGCCCGGCAGGAAGATGGGC
TCCCGTGGACAGGGACTCTTGCTGGCGTACTGCCTGCTCCTTGCCCTTGGCTCTGGCCTGGT
CCTGAGTCGTGTGCCCCATGTCCAGGGGGAACAGCAGGAGTGGGAGGGGACTGAGGAGCTGC
CGTCGCCTCCGGACCATGCCGAGAGGGGCTGAAGAACAACATGAAAAATACAGGCCCAGTCAG
GACCAGGGGCTCCCTGCTTCCCGGTGCTTGCCTGTGTGACCCCGGTACCTCCATGTATCCC
GGCGACCGCGTGCCTCAGATCAACATCACTATCTTGAAAGGGGAGAAGGGTGACCGGGAG
ATCGAGGCTCCAAGGGAATATGGCAAAACAGGCTCAGCAGGGGGCCAGGGGCCACACTGGA
CCCAAGGGCAGAAGGGCTCCATGGGGGCCCTGGGGAGCGGTGCAAGAGCCACTACGCCGC
CTTTTCGGTGGGCCGGAAGAAGCCATGCACAGCAACCACTACTACCAGACGGTGATCTTCG
ACACGGAGTTCTGTGAACCTCTACGACCACTTCAACATGTTACCGGCAAGTTCTACTGCTAC
GTGCCCCGGCCTCTACTTCTTTCAGCCTCAACGTGCACACCTGGAACCAGAAGGAGACCTACCT
GCACATCATGAAGAACGAGGAGGAGGTGGTGATCTTGTTCGCGCAGGTGGGCGACCGCAGCA
TCATGCAAAGCCAGAGCCTGATGCTGGAGCTGCGAGAGCAGGACCAGGTGTGGGTACGCCTC
TACAAGGGCGAACGTGAGAACGCCATCTTCAGCGAGGAGCTGGACACCTACATCACCTTCAG
TGGCTACCTGGTCAAGCACGCCACCGAGCCCCAGTGGCCGGCCACCTCCTTTCTCTCGCC
ACCTTCCACCCCTGCGCTGTGCTGACCCCAACCGCTCTTCCCGATCCCTGGACTCCGACTC
CCTGGCTTTGGCATTCAAGTGAAGACGCCCTGCACACACAGAAAGCCAAAGCGATCGGTGCTCC
CAGATCCCGCAGCCTCTGGAGAGAGCTGACGGCAGATGAAATCACCAGGGCGGGGCACCCGC
GAGAACCCTCTGGGACCTTCCGCGGCCCTCTCTGCACACATCCTCAAGTGACCCCGCACGGC
GAGACGCGGGTGGCGGCAGGGCGTCCCAGGGTGCGGCACCGCGGCTCCAGTCTTTGGAATA
ATTAGGCAAATTTCTAAAGGTCTCAAAGGAGCAAAGTAAACCGTGGAGGACAAAGAAAAGGG
TTGTTATTTTTGTCTTTCCAGCCAGCCTGCTGGCTCCCAAGAGAGAGGCCTTTTCAGTTGAG
ACTCTGCTTAAGAGAAGATCCAAAGTTAAAGCTCTGGGGTCAGGGGAGGGGGCCGGGGCAGG
AAACTACCTCTGGCTTAATTCTTTTAAGCCACGTAGGAACCTTTCTTGAGGGATAGGTGGACC
CTGACATCCCTGTGGCCTTGCCCAAGGGCTCTGCTGGTCTTTCTGAGTCACAGCTGCGAGGT
GATGGGGGCTGGGGCCCCAGGCGTCAGCCTCCCAGAGGGACAGCTGAGCCCCCTGCCTTGGC
TCCAGGTTGGTAGAAGCAGCCGAAGGGCTCCTGACAGTGGCCAGGGACCCCTGGGTCCCCCA
GGCCTGCAGATGTTTCTATGAGGGGCAGAGCTCCTTGGTACATCCATGTGTGGCTCTGCTCC
ACCCCTGTGCCACCCCAAGAGCCCTGGGGGGTGGTCTCCATGCCTGCCACCCTGGCATCGGCT
TTCTGTGCCGCCTCCACACAAATCAGCCCCAGAAGGCCCGGGGCTTGGCTTCTGTTTTT
TATAAAACACCTCAAGCAGCACTGCAGTCTCCATCTCCTCGTGGGCTAAGCATCACCGCTT
CCACGTGTGTTGTGTTGGTTGGCAGCAAGGCTGATCCAGACCCCTTCTGCCCCACTGCCCT
CATCCAGGCCTCTGACCAGTAGCCTGAGAGGGGCTTTTTCTAGGCTTCAGAGCAGGGGAGAG
CTGGAAGGGGCTAGAAAGCTCCCGCTTGTCTGTTTCTCAGGCTCCTGTGAGCCTCAGTCCTG
AGACCAGAGTCAAGAGGAAGTACACGTCCCAATCACCCGTGTCAGGATTCAGTCTCAGGAGC
TGGGTGGCAGGAGAGGCAATAGCCCCCTGTGGCAATTGCAGGACCAGCTGGAGCAGGGTTGCG
GTGTCTCCACGGTGTCTCGCCCTGCCCATGGCCACCCAGACTCTGATCTCCAGGAACCC
ATAGCCCCCTCTCCACCTCACCCATGTTGATGCCCAGGGTCACTCTTGCTACCCGCTGGGCC
CCCAAACCCCGCTGCCTCTCTTCCCTTCCCCCATCCCCACCTGGTTTTGAATAATCCTGC
TTCCCTCTCTGGGCCTGGCTGCCGGGATCTGGGGTCCCTAAGTCCCTCTCTTTAAAGAACTT
CTGCGGGTCAAGTCTGAAGCCGAGTTGCTGTGGGCGTGGCCGGAAGCAGAGCGCCACACTC
GCTGCTTAAGCTCCCCCAGCTCTTTCCAGAAAACATTAAACTCAGAATTGTGTTTTCAA

38/75

FIGURE 38

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41234
><subunit 1 of 1, 281 aa, 1 stop
><MW: 31743, pI: 6.83, NX(S/T): 1
MGSRGQGLLLAYCLLLAFASGLVLSRVPHVQGEQQEWEGTEELPSPPDHAERAEEQHEKYRP
SQDQGLPASRCLRCCDPGTSMPATAVPQINITILKGEKGDRGDRGLQGKYGKTGSAGARGH
TGPKGQKGSMGAPGERCKSHYAAFSVGRKKPMHSNHYYQTVIFDTEFVNLYDHFNMFTGKFY
CYVPGLYFFSLNVHTWNQKETYLHIMKNEEEVVILFAQVGDRSIMQSQSLMLELREQDQVWV
RLYKGERENAI FSEELDTYITFSGYLVKHATEP
```

Signal sequence.

amino acids 1-25

N-glycosylation site.

amino acids 93-97

N-myristoylation sites.

amino acids 7-13, 21-27, 67-73, 117-123, 129-135

Amidation site.

amino acids 150-154

Cell attachment sequence.

amino acids 104-107

FIGURE 39

GAATTCGGCACGAGGGAAGAAGAGAAAGAAAATCTCCGGGGCTGCTGGGAGCATATAAAGAA
GCCCTGTGGCCTTGCTGGTTTTTACCATCCAGACCAGAGTCAGGCCACAGACGGAC**ATGG**CTG
CTCAAGGCTGGTCCATGCTCCTGCTGGCTGTCCTTAACCTAGGCATCTTCGTCCGTCCCTGT
GACACTCAAGAGCTACGATGTCTGTGTATTTCAGGAACACTCTGAATTCATTCCTCTCAAAC
CATTAAAAATATAATGGTGATATTCGAGACCATTTACTGCAACAGAAAGGAAGTGATAGCAG
TCCCAAAAAATGGGAGTATGATTTGTTTGGATCCTGATGCTCCATGGGTGAAGGCTACTGTT
GGCCAATTACTAACAGGTTCCCTACCTGAGGACCTCAAACAAAAGGAATTTCCACCGGCAAT
GAAGCTTCTGTATAGTGTTGAGCATGAAAAGCCTCTATATCTTTCATTTGGGAGACCTGAGA
ACAAGAGAATATTTCCCTTTCCAATTCGGGAGACCTCTAGACACTTTGCTGATTTAGCTCAC
AACAGTGATAGGAATTTTCTACGGGACTCCAGTGAAGTCAGCTTGACAGGCAGTGATGCC**TA**
AAAGCCACTCATGAGGCAAAGAGTTTCAAGGAAGCTCTCCTCCTGGAGTTTTGGCGTTCTCA
TTCTTATACTCTATTCCCGCGTTAGTCTGGTGTATGGATCTATGAGCTCTCTTTTAATATTT
TATTATAAATGTTTTATTTACTTAACTTCCTAGTGAATGTTACAGGTGACTGCTCCCCCAT
CCCCATTTCTTGATATTACATATAATGGCATCATATAACCCCTTTATTGACTGACAAACTACT
CAGATTGCTTAACATTTTGTGCTTCAAAGTCTTATCCCACTCCACTATGGGCTGTTACAGAG
TGCATCTCGGTGTAGAGCAAGGCTCCTTGTCTTCAGTGCCCCAGGGTGAAATACTTCTTTGA
AAAATTTTCATTCATCAGAAAATCTGAAATAAAAAATATGTCTTAATTGAG

40/75

FIGURE 40

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73838
><subunit 1 of 1, 167 aa, 1 stop
><MW: 19091, pI: 7.48, NX(S/T): 1
MAAQGWSMLLLAVLNLGIFVRPCDTQELRCLCIQEHSEFIPLKLIK NIMVIFETIYCNRKEV
IAVPKNGSMICLDPDAPWVKATVGPITNRFLPEDLKQKEFP PAMKLLYSVEHEKPLYLSFGR
PENKRIFPFPPIRETSRHFADLAHNSDRNFLRDSSEVSLTGSDA
```

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 68-72

N-myristoylation site.

amino acids 69-75

Small cytokines (intercrine/chemokine) C-x-C subfamily signature

amino acids 40-85

FIGURE 41

CAGACATGGCTCAGTCACTGGCTCTGAGCCTCCTTATCCTGGTTCTGGCCTTTGGCATCCCC
AGGACCCAAGGCAGTGATGGAGGGGCTCAGGACTGTTGCCTCAAGTACAGCCAAAGGAAGAT
TCCCGCCAAGGTTGTCCGCAGCTACCGGAAGCAGGAACCAAGCTTAGGCTGCTCCATCCCAG
CTATCCTGTTCTTGCCCCGCAAGCGCTCTCAGGCAGAGCTATGTGCAGACCCAAAGGAGCTC
TGGGTGCAGCAGCTGATGCAGCATCTGGACAAGACACCATCCCCACAGAAACCAGCCCAGGG
CTGCAGGAAGGACAGGGGGGCCTCCAAGACTGGCAAGAAAGGAAAGGGCTCCAAGGCTGCA
AGAGGACTGAGCGGTCACAGACCCCTAAAGGGCCATAGCCCAGTGAGCAGCCTGGAGCCCTG
GAGACCCACACAGCCTCACCAGCGCTTGAAGCCTGAACCCAAGATGCAAGAAGGAGGCTATG
CTCAGGGGGCCCTGGAGCAGCCACCCCATGCTGGCCTTGCCACACTCTTTCTCCTGCTTTAAC
CACCCCATCTGCATTCCCAGCTCTACCCTGCATGGCTGAGCTGCCCACAGCAGGCCAGGTCC
AGAGAGACCGAGGAGGGAGAGTCTCCCAGGGAGCATGAGAGGAGGCAGCAGGACTGTCCCCT
TGAAGGAGAATCATCAGGACCCTGGACCTGATACGGCTCCCCAGTACACCCACCTCTTCCT
TGTAATATGATTTATACCTAACTGAATAAAAAGCTGTTCTGTCTTCCCNCCCC

42/75

FIGURE 42

><MW: 14646, pI: 10.45, NX(S/T): 0

MAQSLALSLILVLAFGIPTQGSDDGAQDCCLKYSQRKIPAKVVRSYRKQEPSLGCSIPAI
LFLPRKRSQAELCADPKELWVQQLMQHLDKTPSPQKPAQGCRCRKGASKTGKKGKSGSKGCKR
TERSQTPKGP

Important features of the protein:

Signal peptide:

amino acids 1-17

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 67-71

N-myristoylation sites.

amino acids 17-23, 23-29, 27-33, 108-114, 118-124, 121-127

Amidation site.

amino acids 112-116

Small cytokines.

amino acids 51-91

FIGURE 43

AAGGAGCAGCCCGCAAGCACCAAGTGAGAGGCATGAAGTTACAGTGTGTTTCCCTTTGGCTC
CTGGGTACAATACTGATATTGTGCTCAGTAGACAACCACGGTCTCAGGAGATGTCTGATTTT
CACAGACATGCACCATATAGAAGAGAGTTTCCAAGAAATCAAAAGAGCCATCCAAGCTAAGG
ACACCTTCCCAAATGTCACCTATCCTGTCCACATTGGAGACTCTGCAGATCATTAAGCCCTTA
GATGTGTGCTGCGTGACCAAGAACCTCCTGGCGTTCTACGTGGACAGGGTGTTCAAGGATCA
TCAGGAGCCAAACCCCAAATCTTGAGAAAAATCAGCAGCATTGCCAACTCTTTCCTCTACA
TGCAGAAAACCTCTGCGGCAATGTCAGGAACAGAGGCAGTGTCACTGCAGGCAGGAAGCCACC
AATGCCACCAGAGTCATCCATGACAACCTATGATCAGCTGGAGGTCCACGCTGCTGCCATTAA
ATCCCTGGGAGAGCTCGACGTCTTTCTAGCCTGGATTAATAAGAATCATGAAGTAATGTTCT
CAGCTTGATGACAAGGAACCTGTATAGTGATCCAGGGATGAACACCCCCTGTGCGGTTTACT
GTGGGAGACAGCCACCTTGAAGGGGAAGGAGATGGGGAAGGCCCTTGCAGCTGAAAGTCC
CACTGGCTGGCCTCAGGCTGTCTTATTCCGCTTGAAAATAGGCAAAAAGTCTACTGTGGTAT
TTGTAATAAACTCTATCTGCTGAAAGGGCCTGCAGGCCATCCTGGGAGTAAAGGGCTGCCTT
CCCATCTAATTTATTGTAAAGTCATATAGTCCATGTCTGTGATGTGAGCCAAGTGATATCCT
GTAGTACACATTGTACTGAGTGGTTTTTCTGAATAAATTCCATATTTTACCTATGA

FIGURE 44

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282
><subunit 1 of 1, 177 aa, 1 stop
><MW: 20452, pI: 8.00, NX(S/T): 2
MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST
LETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMOKTLRQCQEQ
RQCHCRQEATNATRVIHDNYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA
```

Signal sequence:

amino acids 1-18

N-glycosylation sites.

amino acids 56-60, 135-139

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation site.

amino acids 24-30

Actinin-type actin-binding domain signature 1.

amino acids 159-169

45/75

FIGURE 45

GCTCCCAGCCAAGAACCTCGGGGCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGCCG
CTAAGCGAGGCCTCCTCCTCCCGCAGATCCGAACGGCCTGGGCGGGGTACCCCGGCTGGGA
CAAGAAGCCGCCGCTGCCTGCCCGGGCCCGGGGAGGGGGCTGGGGCTGGGGCCGGAGGCGG
GGTGTGAGTGGGTGTGTGCGGGGGGCGGAGGCTTGATGCAATCCCGATAAGAAATGCTCGGG
TGTCTTGGGCACCTACCCGTGGGGCCCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCCGAG
CCGCCGCGCCGTAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCAACCC
GGCACTCACAGCCCCGAGCGCATCCCGGTGCGCGCCAGCCTCCCGCACCCCCATCGCCGG
AGCTGCGCCGAGAGCCCCAGGGAGGTGCC**ATG**CGGAGCGGGTGTGTGGTGGTCCACGTATGG
ATCCTGGCCGGCCTCTGGCTGGCCGTGGCCGGGGCGCCCCCTCGCCTTCTCGGACGCGGGGCC
CCACGTGCACTACGGCTGGGGCGACCCCATCCGCCTGCGGCACCTGTACACCTCCGGCCCCC
ACGGGCTCTCCAGCTGCTTCTGCGCATCCGTGCCGACGGCGTCGTGGACTGCGCGCGGGGC
CAGAGCGCGCACAGTTTGCTGGAGATCAAGGCAGTCGCTCTGCGGACCGTGGCCATCAAGGG
CGTGACAGCGTGCGGTACCTCTGCATGGGCGCCGACGGCAAGATGCAGGGGGCTGCTTCAGT
ACTCGGAGGAAGACTGTGCTTTTCGAGGAGGAGATCCGCCCAGATGGCTACAATGTGTACCGA
TCCGAGAAGCACCGCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAA
CAGAGGCTTTCTTCCACTCTCTCATTTTCTGCCCCATGCTGCCCATGGTCCCAGAGGAGCCTG
AGGACCTCAGGGGCCACTTGGAATCTGACATGTTCTCTTCGCCCCCTGGAGACCGACAGCATG
GACCCATTTGGGCTTGTCACCGGACTGGAGGCCGTGAGGAGTCCAGCTTTGAGAAG**TAA**CT
GAGACCATGCCCGGGCCTCTTCACTGCTGCCAGGGGGCTGTGGTACCTGCAGCGTGGGGGACG
TGCTTCTACAAGAACAGTCCTGAGTCCACGTTCTGTTTCTAGCTTTAGGAAGAAACATCTAGAA
GTTGTACATATTCAGAGTTTTCCATTGGCAGTGCCAGTTTCTAGCCAATAGACTTGTCTGAT
CATAACATTGTAAGCCTGTAGCTTGCCCAGCTGCTGCCTGGGCCCCCATTTCTGCTCCCTCGA
GGTTGCTGGACAAGCTGCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAAC
TCACTTCCTTTGGAAAAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTTCTCATCACTTC
CCCAGGAGCAGCCAGAAGACAGGCAGTAGTTTTTAATTTAGGAACAGGTGATCCACTCTGTA
AAACAGCAGGTAAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG
ACCATTTGCCCTTCCCAAATCCCTCCAGGCCAGAACTGACTGGAGCAGGCATGGCCCACCAG
GCTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACCTGAGAATTCCCC
CTGAGGCCAGTTCTGTCATGGATGCTGTCTGAGATAAATTGCTGTCCCGGTGTACCTGC
TTCCATCTCCAGCCCACCAGCCCTCTGCCACCTCACATGCCTCCCCATGGATTGGGGCCT
CCCAGGCCCCCCCCACCTTATGTCAACCTGCACTTCTTGTTCAAAAATCAGGAAAAGAAAAGAT
TTGAAGACCCCAAGTCTTGTCAATAACTTGCTGTGTGGAAGCAGCGGGGGAAGACCTAGAAC
CCTTTCCCCAGCACTTGGTTTTCCAACATGATATTTATGAGTAATTTATTTTGATATGTACA
TCTCTTATTTTCTTACATTATTTATGCCCCCAAATTATTTATGTATGTAAGTGAGGTTTG
TTTTGTATATTAAATGGAGTTTGTGTGT

FIGURE 46

MRSGCVVHVWILAGLWLAVAGRPLAFSDAGPHVHYGWGDPIRLRHLYTSGPHGLSSCFLRI
RADGVVDCARGQSAHSLLEIKAVALRTVAIKGVHVSRYLCMGADGKMQGLLQYSEEDCAFEE
EIRPDGYNVYRSEKHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLRGHLESD
MFSSPLETDSMDPFGLVTGLEAVRSPSF EK

Signal peptide:

amino acids 1-22

Casein kinase II phosphorylation site.

amino acids 78-82, 116-120, 190-194, 204-208

N-myristoylation site.

amino acids 15-21, 54-60, 66-72, 201-207

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 48-59

FIGURE 47

GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTCAGTGGCCTGATCGCGATGGGGACAAA
GGCGCAAGTCGAGAGGAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG
CATTGGGCAGTGTTACAGTGCACCTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT
GTGAAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAGTTTGACCA
AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG
TGACCTTCTTGCCAACCTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC
ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT
GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCCTCCTCTGCCACCATTGGGAACCGGG
CAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGAT
GGGATAGTGATGCCTACGAATCCCAAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCTT
GAATCCCACAACAGGAGAGCTGGTCTTTGATCCCCCTGTCAGCCTCTGATACTGGAGAATACA
GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT
GTGGAGCGGAATGTGGGGGTCTCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAAT
CTTGGTTTTTTGGCATCTGGTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGA
CTTCGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG
ACCTCGTCATTCCTGGTTGAGCCTGGTCGGCTCACCGCCTATCATCTGCATTTGCCTTACT
CAGGTGCTACCGGACTCTGGCCCCTGATGTCTGTAGTTTCACAGGATGCCTTATTTGTCTTC
TACACCCACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC
ATCCTCCTTCATGCCCTCCCTCCCTTTCTTACCCTGCTGAGTGGCCTGGAACCTGTTTAAA
GTGTTTATTCCCCATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC
TTCTAAGTAGACAGCAAAAATGGCGGGGGTTCGAGGAATCTGCACTCAACTGCCCACCTGGC
TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGGTTCTGGGCTCTTTCCTTGTGTACTGAC
GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG
TGATGACACTGGGGTCCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG
CCACTGGGATCCCTCTGCCCTGTCCTCCTGAATACAAGCTGACTGACATTGACTGTGTCTGT
GGAAAATGGGAGCTCTTGTTGTGGAGAGCATAGTAAATTTTCAGAGAAGTTGAAGCCAAAAG
GATTTAAAACCGCTGCTCTAAAGAAAAGAAAAGTGGAGGCTGGGCGCAGTGGCTCACGCCTG
TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTGCGGAGTTTCGGGATCAGCCTGACCA
ACATGGAGAAACCCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTC
CCAGCTGCTCAGGAGCCTGGCAACAAGAGCAAACTCCAGCTCAAAAAAAAAAAAAAAAAA

FIGURE 48

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPRVEW
KFDQGD TTRLVCYNNKITASYEDRVTF LPTGITFKSVTREDTGTYTCMVSEEGGNSYGEVKV
KLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPSEYTWFKDGIVMPTNPKSTRAFSNS
SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGT PMTSNAVRMEAVERNVGVIVA AVLVTLI
LLGILVFGIWFAYS SRGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 238-255

N-glycosylation site.

amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 270-274

Casein kinase II phosphorylation site.

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158,
193-197, 203-207, 287-291

N-myristoylation site.

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

CCCACGCGTCCGAACCTCTCCAGCG**ATG**GGGAGCCGCCCGCCTGCTGCCCAACCTCACTCTGT
GCTTACAGCTGCTGATTCTCTGCTGTCAAACCTCAGTACGTGAGGGACCAGGGCGCCATGACC
GACCAGCTGAGCAGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGAAGCA
CGTGACAGGTCACCGGGCGTCGCATCTCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGC
TCATAGTGAGACGGACACGTTTGGCAGCCGGGTTCGCATCAAAGGGGGCTGAGAGTGAGAAG
TACATCTGTATGAACAAGAGGGGGCAAGCTCATCGGGAAGCCCAGCGGGAAGAGCAAAGACTG
CGTGTTACGGAGATCGTGCTGGAGAACAACATACGGCCTTCCAGAACGCCCGGCACGAGG
GCTGGTTTCATGGCCTTCACGCGGCAGGGGCGGCCCCGCCAGGCTTCCCGCAGCCGCCAGAAC
CAGCGCGAGGGCCCACTTCATCAAGCGCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCACGC
CGAGAAGCAGAAGCAGTTCGAGTTTGTGGGCTCCGCCCCCACC CGCCGGACCAAGCGCACAC
GGCGGCCCCAGCCCCTCACG**TAG**TCTGGGAGGCAGGGGGGCAGCAGCCCCTGGGGCGCCTCCC
CACCCCTTTCCCTTCTTAATCCAAGGACTGGGCTGGGGTGGCGGGAGGGGAGCCAGATCCCC
GAGGGAGGACCCTGAGGGCCGCGAAGCATCCGAGCCCCAGCTGGGAAGGGGCAGGCCGGTG
CCCCAGGGGCGGCTGGCACAGTGCCCCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAACT
GAGTGTCACCCTGATCTCAGGCCACCAGCCTCTGCCGGCCTCCAGCCGGGCTCCTGAAGCC
CGCTGAAAGGTCAGCGACTGAAGGCCTTGCAGACAACCGTCTGGAGGTGGCTGTCTCTCAAAA
TCTGCTTCTCGGATCTCCCTCAGTCTGCCCCAGCCCCCAAACCTCCTCCTGGCTAGACTGTA
GGAAGGGACTTTTGT TTGTTTGT TTGTTTGT TTCAGGAAAAAAGAAAGGGAGAGAGAGGAAAAATAG
AGGGTTGTCCACTCCTCACATTCCACGACCCAGGCCTGCACCCCACCCCCAACTCCAGCCC
CGGAATAAAAACCATTTTTCTGTC

FIGURE 50

MGAARLLPNLTLCLQLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRISATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLGKPSGKSKDCVFTEIVLENNYTAFQONARHEGWEMAFTRQGRPRQASRSRQNQREAHFIKRLYQGQLPFPNHAEKQKQFEFVGSAPTRRTKRTRRPQPLT

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 60-64

Casein kinase II phosphorylation site.

amino acids 65-69

Tyrosine kinase phosphorylation site.

amino acids 39-48, 89-97

N-myristoylation site.

amino acids 69-75, 188-194

Amidation site.

amino acids 58-62

HBGF/FGF family signature.

amino acids 103-128

FIGURE 51

GTTGTGTCCTTCAGCAAAACAGTGGATTTAAATCTCCTTGCACAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAGAAGAAAAAGAAGAAG
AAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC
GGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATT
GACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTTGTAGAGATTTT
TTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCCAAAGCGGTTGGCTTTGTGAGTGAAGAC
GAATACTTGGAATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAAGGTCACCGTGAACCTATCCACCATA
TTTCAGAAGCCAAGGGTACAGGTGTCCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCCTCAGCAGAATTCAGTGGTACAAGGATGACAAAAGACTGATTTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTTCCTCTCAAACTCATCTTCTTCAATGTCTCTG
AACATGACTATGGGAACTACACTTGGCTGGCCTCCAACAAGCTGGGCCACACCAATGCCAGC
ATCATGCTATTTGGTCCAGGCGCCGTCAGCGAGGTGAGCAACGGCACGTCGAGGAGGGCAGG
CTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGACCTGCTTCTCAAATTTTGATGTGAGTGCC
ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCAACACAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGGAAAAGAGTTTTAAAAAAGAAATTGAA
AATTGCCTTGCAGATATTTAGGTACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGAACCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA
GGGCTCAGCCTCTCTGCCCACAGAGTGCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAAGAGCAAAA
AAAAA

52/75

FIGURE 52

MKTIQPKMHNSISWAI FTGLAALCLFQGV PVRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLT CIATGRPEPTVTWRHISPKAVGFVSEDEYL
EIQGITREQSGDYEC SASNDVAAPVVR RVKVTVNYPPYISEAKGTGVPVGQKGT LQCEASAV
PSAEFQWYKDDKRLIEGKKG VKVENRPFLSKLI FFNVSEHDYGN YTCVASNKLGH TNASIML
FGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF

Signal peptide:

amino acids 1-28

[illegible]

54/75

FIGURE 54

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59212

><subunit 1 of 1, 440 aa, 1 stop

><MW: 42208, pI: 6.36, NX(S/T): 1

MKFQGPLACLLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGA
AGSKVSEALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHG
ADAVRGSWQGVPGHSGAWETSGGHGIFGSQGGGLGGQGGNPGGLGTPWVHGYPGNSAGSFGM
NPQGAPWGQGGNGGPPNFGTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGS
GSQSGSSGSGSNGDNNNGSSSGSSSGSSSGSSSGSSSGSSSGSSSGSSSGNSGSSRGDSGSESSW
GSSTGSSSGNHGSGGGNGHKPGCEKPGNEARGSGESGIQGFRGQGVSSNMREISKEGNRLL
GGSGDNYRGQGSWSGGGDAVGGVNTVNSETSPGMFNFDTFWKNFKSKLGFINWDAINKDQ
RSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70,
74-80, 90-96, 96-102, 130-136, 140-146, 149-155, 152-158,
155-161, 159-165, 163-169, 178-184, 190-196, 194-200, 199-205,
218-224, 236-242, 238-244, 239-245, 240-246, 245-251, 246-252,
249-252, 253-259, 256-262, 266-272, 270-276, 271-277, 275-281,
279-285, 283-289, 284-290, 287-293, 288-294, 291-297, 292-298,
295-301, 298-304, 305-311, 311-317, 315-321, 319-325, 322-328,
323-329, 325-331, 343-349, 354-360, 356-362, 374-380, 381-387,
383-389, 387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

55/75

FIGURE 55

AGCCAGGCAGCACATCACAGCGGGAGGAGCTGTCCCAGGTGGCCCAGCTCAGCA**AT**GGCAAT
GGGGGTCCCCAGAGTCATTCTGCTCTGCCTCTTTGGGGCTGCGCTCTGCCTGACAGGGTCCC
AAGCCCTGCAGTGCTACAGCTTTGAGCACACCTACTTTGGCCCCCTTTGACCTCAGGGCCATG
AAGCTGCCCAGCATCTCCTGTCCTCATGAGTGCTTTGAGGGCTATCCTGTCTCTGGACACCGG
GTATCGCGCGCCGGTGACCCTGGTGCGGAAGGGCTGCTGGACCGGGCCTCCTGCGGGCCAGA
CGCAATCGAACCCGGACGCGCTGCCGCCAGACTACTCGGTGGTGCGCGGCTGCACAACCTGAC
AAATGCAACGCCCCACCTCATGACTCATGACGCCCTCCCCAACCTGAGCCAAGCACCCGACCC
GCCGACGCTCAGCGGCGCCGAGTGCTACGCCTGTATCGGGGTCCACCAGGATGACTGCGCTA
TCGGCAGGTCCCGACGAGTCCAGTGTCACCAGGACCAGACCGCCTGCTTCCAGGGCAGTGGC
AGAATGACAGTTGGCAATTTCTCAGTCCCTGTGTACATCAGAACCTGCCACCGGCCCTCCTG
CACCACCGAGGGCACCACCAGCCCCCTGGACAGCCATCGACCTCCAGGGCTCCTGCTGTGAGG
GGTACCTCTGCAACAGGAAATCCATGACCCAGCCCTTACCAGTGCTTACAGCCACCACCCCT
CCCCGAGCACTACAGGTCCTGGCCCTGCTCCTCCCAGTCCTCCTGCTGGTGGGGCTCTCAGC
A**TA**GACCGCCCCCTCCAGGATGCTGGGGACAGGGCTCACACACCTCATTTCTTGCTGCTTCAGC
CCCTATCACATAGCTCACTGGAAAATGATGTTAAAGTAAGAATTGCAAAA

56/75

FIGURE 56

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA86576
><subunit 1 of 1, 251 aa, 1 stop
><MW: 26935, pI: 7.42, NX(S/T): 2
MAMGVPRVILLCLFGAALCLTGSQALQCYSFEHTYFGPFDLRAMKLPSISCPHECFEAILSL
DTGYRAPVTLVRKGCWTGPPAGQTQSNPDALPPDYSVVRGCTTDKCNAHLMTHDALPNLSQA
PDPPTLSGAECYACIGVHQDDCAIGRSRRVQCHQDQTACFQGSGRMTVGNFSPVYIRTCHR
PSCTTEGTTSPWTAIDLQGSCCEGYLCNRKSMTQPFTSASATTPPRALQVLALLLPVLLLVG
LSA

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 233-251

N-glycosylation sites.

amino acids 120-124, 174-178

N-myristoylation sites.

amino acids 15-21, 84-90

GGAGCCGCCCTGGGTGTGTCAGCGGCTCGGCTCCCGCGCACAGCTCCGGCCGTGCGCGCAGCCCTCG
GCACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCCCTGACTCCGTCCCGGCCAGGGAGGGC
CATGATTTCCCTCCCGGGGCCCTGGTGACCAACTTGCTGCGGTTTTTGTTCCTGGGGCTGA
GTGCCCTCGCGCCCCCTCGCGGGGCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAG
GCGGTGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGACGGGGAGGTGTCTTC
ATCCCAGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATC
AGGTGTTGTCCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC
ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA
CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT
TAGAACTCAATGTACTGGTTCCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT
GTGGGGGCAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA
GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG
GGTCTTTAAGCCTCACCAACCTTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCAC
AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGC
AGTGGTTGCTGGAGCTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCC
TCTTGTACCACCGCCGGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCC
ATTGCTCCCCGGACCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT
TTCCTCTGTACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCAGGCCTGGTGCAT
TGACCCCCACGCCAGTCTCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCACGACAGAT
GGGGCCCACCCTCAACCAATATCCCCCATCCCTGGTGGGGTTTCTTCCTCTGGCTTGAGCCG
CATGGGTGCTGTGCCTGTGATGGTGCCTGCCAGAGTCAAGCTGGCTCTCTGGTAT**TGA**TGAC
CCCACCACTCATTTGGCTAAAGGATTTGGGGTCTCTCCTTCCCTATAAGGGTCACCTCTAGCAC
AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCACCTCTC
TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA
AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTG
CTGAAATTAGCTACTCACCAAGAGTGAGGGGCAGAGACTTCCAGTCACTGAGTCTCCCAGGC
CCCCTTGATCTGTACCCACCCCTATCTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGT
ATTGATATAACCTGTCAGGCTGGCTTGGTTAGGTTTTTACTGGGGCAGAGGATAGGGAATCTC
TTATTAAACTAACATGAAATATGTGTTGTTTTTCATTTGCAAATTTAAATAAAGATACATAA
TGTTTGTATGAAAAA

FIGURE 58

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSS
SQPWEVPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKGDSGPY
SCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLOQVPHVGANVTLSQCSPRSKPAVQYQ
WDRQLPSFQTFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQC�NVTLLEVSTGPGAA
VVAGAVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLWPWKSSDTISKNGTL
SSVTSARALRPPHGP RP RGALTP T PSLSSQALPSPRLPTTDGAHPQPISPIPGGVSSSSGLSR
MGAVPVMVPAQSQAGSLV

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 245-267

N-glycosylation site.

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

N-myristoylation site.

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262,
262-268, 308-314, 363-369, 364-370

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 164-175

FIGURE 59

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTTTGCACATGGAG
 GACAGCAGCAAAGAGGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC
 CATA CGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGT
 CATTTTGATTTTGCTGTTTATTTTTTTTTTCTTTTTCTTTTTCCCACCACATTGTATTTTAT
 TTCCGTACTTCAGAA**ATG**GGCCTACAGACCACAAAGTGGCCCAGCCATGGGGCTTTTTTCCT
 GAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAAACTCCTGGCCTGCC
 CTAGTGTGTGCCGCTGCGACAGGAACCTTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG
 CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGC
 TGGATTTCCTGCAGAACTGCACAATGTACAGTCGGTGCACACGGTCTACCTGTATGGCAACC
 AACTGGACGAATTCCCCATGAACCTTCCCAAGAATGTACAGAGTTCTCCATTTGCAGGAAAAC
 AATATTACAGACCATTTACGGGCTGCTCTTGCCAGCTCTTGAAGCTGAAGAGCTGACCCCT
 GGATGACAACTCCATATCCACAGTGGGGTGGGAAGACGGGGCCTTCCGGGAGGCTATTAGCC
 TCAAATTGTTGTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCCTGTGGAC
 TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTCATATCCGACATGGCCTTCCAGAA
 TCTCACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCG
 AGGGCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCC
 CACCCTCCTCCCGATCTCCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACCAGAT
 AAACCACATTCCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA
 ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTC
 ACTGCTCGGAATAACCCTTGGTTTTGTGACTGCAGTATTAAATGGGTACAGAATGGCTCAA
 ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCCAAGGTCCTGAACAAGTCCGGG
 GGATGGCCGTCAGGGAATTAAATATGAATCTTTTGTCTGTCCCACCACGACCCCCGGCCTG
 CCTCTCTTCACCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTCTAT
 TCCAAACCCTAGCAGAAGCTACACGCCTCCAACCTCCTACCACATCGAAACTTCCCACGATTC
 CTGACTGGGATGGCAGAGAAAGAGTGACCCACCTATTTCTGAACGGATCCAGCTCTCTATC
 CATTTTGTGAATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTCTTACCCGTGATGGCATA
 CAACTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATCGTTCAGGAGCGCATAG
 TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATT
 TGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTTAGAAGACACCATTTGTTTCAGAGGC
 CACCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA
 CGTCCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATTT
 GTGCTGGTGGTCTTGCTCAGCGTCTTTTGTGCTGGCATATGCACAAAAAGGGGCGCTACACCTC
 CCAGAAGTGGAATAACAACCGGGGGCCGGCGGAAAGATGATTAT'TGCGAGGCAGGCACCAAGA
 AGGACAACCTCCATCCTGGAGATGACAGAAACCAGTTTTTCAGATCGTCTCCTTAAATAACGAT
 CAACTCCTTAAAGGAGATTTACAGACTGCAGCCCATTTACACCCCAAATGGGGGCATTAATTA
 CACAGACTGCCATATCCCCAACAACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC
 ACTGCCATACG**TGA**CAGCCAGAGGGCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAA
 CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT
 TTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAGTG
 CTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTGTAACTCTTTGCTTTTTAAATCTT

FIGURE 60

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIP
 EGVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTI
 SRAALAQLLKLEELHLDDNSISTVGVEDGAFREAI SLKLLFLSKNHLSSVPVGLPVDLQELR
 VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD
 LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN
 PWFCDCSIKWWTEWLKYIPSSLNVRGFMCGPEQVRGMAVRELMNLLSCPTTTPGLPLFTP
 APSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLP TPDWDGRERVTPPISERIQLSIHFVND
 TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPL
 DAFNYRAVEDTICSEATTHASYLNNGSNTASSHEQTTSMSGSPFLAGLIGGAVIFVLVVL
 LSVFCWHMHKKGRYTSQKWKYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKG
 DFR LQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 542-561

N-glycosylation site.

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

Casein kinase II phosphorylation site.

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

Tyrosine kinase phosphorylation site.

amino acids 319-328

N-myristoylation site.

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,
 522-528, 545-551, 633-639

Amidation site.

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

61/75

FIGURE 62

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906
><subunit 1 of 1, 772 aa, 1 stop
><MW: 87002, pI: 4.64, NX(S/T): 8
MNCYLLLRFMLGIPLLWPCLGATENSQTKKVKQPVRSHLRVKRGVWVNQFFVPEEMNTTSHH
IGQLRSDLDNGNNSFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVSDINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARL
LYSLAQGPYFSVEPTTGIVIRISSKMDRELQDEYWVIIQAKDMIGQPGALSGTTSVLIKLSD
VNDNKPIFKESLYRLTVSESAPTGTSGTIMAYDNDIGENAEMDYSIEEDDSQTFDIITNHE
TQEGIVILKKKVD FEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFL
PYYVFEVFEETPQGSFVGVSATDPDNRKSPIRYSITRSKVFENINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTI SAVDR
DESIEEHHFYFNLSVEDTNSSFTIIDNQDNTAVILTNRTGFNLQEEPVFYISILIADNGIP
SLTSTNTLTIHVCDGDSGSTQTCQYQELVLSMGFKTEVIIAILICIMIIFGFIFLTLGLKQ
RRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKT TSAEIRSLY
RQSLQVGPD SAIFRKFILEKLEEANTDPCAPPFDSLQTYAFEGTGS LAGSLSSLES AVSDQD
ESYDYLNELGPRFKRLACMFGSAVQSNN
```

Important features:**Signal peptide:**

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518,
516-519 and 534-537**Cadherins extracellular repeated domain signature.**

amino acids 136-146 and 244-254

63/75

FIGURE 63

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCA**ATGGC**
CGCCCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTC
TCTTGGCCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCAGTGCAGGCTTGAC
AAGTCCAACCTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAG
CTTGGCTGATAACAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTA
TGAGTGAGCGCTGCTATCTGATGAAGCAGGTGCTGAACCTTACCCTTGAAGAAGTGCTGTTCC
CCTCAATCTGATAGGTTCCAGCCTTATATGCAGGAGGTGGTGCCCTTCCTGGCCAGGCTCAG
CAACAGGCTAAGCACATGTCATATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAA
AGCTGAAGGACACAGTGAAAAAGCTTGGAGAGAGTGGAGAGATCAAAGCAATTGGGAGAACTG
GATTTGCTGTTTATGTCTCTGAGAAATGCCTGCATT**TGACC**AGAGCAAAGCTGAAAAATGAA
TAACTAACCCCTTTCCCTGCTAGAAATAACAATTAGATGCCCCAAAGCGATTTTTTTTAAAC
CAAAAGGAAGATGGGAAGCCAAACTCCATCATGATGGGTGGATTCCAAATGAACCCCTGCGT
TAGTTACAAAGGAAACCAATGCCACTTTTGTTTATAAGACCAGAAGGTAGACTTTCTAAGCA
TAGATATTTATTGATAACATTTTCATTGTAAGTGGTGTCTATACACAGAAAACAATTTATTT
TTTAAATAATTGTCTTTTCCATAAAAAAGATTACTTTCCATTCCTTTAGGGGAAAAAACCC
CTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAATGTATTTATTATTATTA
TAAGACTGCATTTTATTTATATCATTTTATTAATATGGATTTATTTATAGAAACATCATTCG
ATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTATAGAGCTATAAC
ATGTTTATTTGACCTCAATAAACACTTGGATATCCC

64/75

FIGURE 64

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA125185
><subunit 1 of 1, 179 aa, 1 stop
><MW: 20011, pI: 8.10, NX(S/T): 3
MAALQKSVSSFLMGTLATSCLLLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKE
ASLADNNTDVRLIGEKLFFHGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVPFLAR
LSNRLSTCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI
```

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 65

GCCCTAACCTTCCCAGGGCTCAGCTCTTTGGAGCTGCCCATTCCTCCGGCTGCGAGAAAGGA
 CGCGCGCCCTGCGTCGGGCGAAGAAAAGAAGCAAAACTTGTTCGGGAGGGTTTCGTCATCAAC
 CTCCTTCCCGCAAACCTAAACCTCCTGCCGGGGCCATCCCTAGACAGAGGAAAGTTCTTGCA
 GAGCCGACCAGCCCTAGTGGATCTGGGGCAGGCAGCGGCGCTGGCTGTGGAATTAGATCTGT
 TTTGAACCCAGTGGAGCGCATCGCTGGGGCTCGGAAGTCACCGTCCGCGGGCACC GGTTGG
 CGCTGCCCCGAGTGGAAACCGACAGTTTTCGAGCCCTCGGCTGCAAGTGGCCTCTCCTCCCCGCG
 GTTGTGTTTCAGTGTTCGGGTGAGGGCTGCGAGTGTGGCAAGTTGCAAAGAGAGCCTCAGAGG
 TCCGAAGAGCGCTGCGCTCCTACTCGCGTTTCGCTTCTTCTCCTCTTCGGTTCCCTACTGTGA
 AATCGCAGCGACATTTACAAAGGCCTCCGGGTCTTACCGAGACCGATCCGCAGCGTTTGGCC
 CGGTCGTGCCTATTGCATCGGGAGCCCCCGAGCACC GGCGAAATGCGGAGGTTCCCCGAAGGC
 CGACCTGGCCGCTGCAGGAGTTATGTTACTTTGCCACTTCTTCACGGACCAGTTTCAGTTTCG
 CCGATGGGAAACCCGGAGACCAAATCCTTGATTGGCAGTATGGAGTTACTCAGGCCCTTCCCT
 CACACAGAGGAGGAGGTGGAAGTTGATTCACACGCGTACAGCCACAGGTGGAAAAGAACTT
 GGACTTTCTCAAGGCGGTAGACACGAACCGAGCAAGCGTCGGCCAAGACTCTCCTGAGCCCA
 GAAGCTTCACAGACCTGCTGCTGGATGATGGGCAGGACAATAACACTCAGATCGAGGAGGAT
 ACAGACCACAATTACTATATATCTCGAATATATGGTCCATCTGATTCTGCCAGCCGGGATTT
 ATGGGTGAACATAGACCAAATGGAAAAAGATAAAGTGAAGATTCATGGAATATTGTCCAATA
 CTCATCGGCAAGCTGCAAGAGTGAATCTGTCCTTCGATTTTCCATTTTATGGCCACTTCCTA
 CGTGAAATCACTGTGGCAACCGGGGGTTTCATATACACTGGAGAAGTCGTACATCGAATGCT
 AACAGCCACACAGTACATAGCACCTTTAATGGCAAATTTTCGATCCCAGTGTATCCAGAAATT
 CAACTGTCAGATATTTTGATAATGGCACAGCACTTGTGGTCCAGTGGGACCATGTACATCTC
 CAGGATAATTATAACCTGGGAAGCTTCACATTCCAGGCAACCCCTGCTCATGGATGGACGAAT
 CATCTTTGGATACAAAGAAATTCCTGTCTTGGTCACACAGATAAGTTCAACCAATCATCCAG
 TGAAAGTCGGACTGTCCGATGCATTTGTCTGTGCCACAGGATCCAACAAATTCCCAATGTT
 CGAAGAAGAACAATTTATGAATACCACCGAGTAGAGCTACAAATGTCAAAAATTACCAACAT
 TTCGGCTGTGGAGATGACCCCATTAACCCACATGCCTCCAGTTTAAACAGATGTGGCCCCCTGTG
 TATCTTCTCAGATTGGCTTCAACTGCAGTTGGTGTAGTAACTTCAAAGATGTTCCAGTGGA
 TTTGATCGTCATCGGCAGGACTGGGTGGACAGTGGATGCCCTGAAGAGTCAAAAGAGAAGAT
 GTGTGAGAATACAGAACCAGTGGAAACTTCTTCTCGAACCACCACAACCGTAGGAGCGACAA
 CCACCCAGTTTCAGGGTCCTAACTACCACCAGAAGAGCAGTGACTTCTCAGTTTCCCACCAGC
 CTCCCTACAGAAGATGATACCAAGATAGCACTACATCTAAAAGATAATGGAGCTTCTACAGA
 TGACAGTGCAGCTGAGAAGAAAGGGGGAACCCCTCCACGCTGGCCTCATCATTGGAATCCTCA
 TCCTGGTCCCTCATTGTAGCCACAGCCATTCTTGTGACAGTCTATATGTATCACCACCCAACA
 TCAGCAGCCAGCATCTTCTTTATTGAGAGACGCCCCAAGCAGATGGCCTGCGATGAAGTTTAG
 AAGAGGCTCTGGACATCCTGCCTATGCTGAAGTTGAACCAGTTGGAGAGAAAGAAGGCTTTA
 TTGTATCAGAGCAGTGCTTAAATTTCTAGGACAGAACAACACCAGTACTGGTTTACAGGTGT
 TAAGACTAAAATTTTGCCTATACCTTTAAGACAAACAAACACACACACAAACAAGCTC
 TAAGCTGCTGTAGCCTGAAGAAGACAAGATTTCTGGACAAGCTCAGCCCAGGAAACAAAGGG
 TAAACAAAAAACTAAACTTATACAAGATACCATTTTACTGAACATAGAATTCCCTAGTGG
 AATGTCATCTATAGTTCACTCGGAACATCTCCCGTGGACTTATCTGAAGTATGACAAGATTA
 TAATGCTTTTGGCTTAGGTGCAGGGTTGCAAAGGGGATCAGAAAAAAAATCATAATAAAGC
 TTTAGTTCATGAGGG

FIGURE 66

MARFPKADLAAAGVMLLCHFFTDQFQFADGKPGDQILDWQYGVTAQFPHTEEVEVD SHAYS
 HRWKRNLDFLKAVDTNRASVGQDSPEPRSFTDLLLDDGQDNNTQIEEDTDHNYIISRIYGPS
 DSASRDLWVNIDQMEKDKVKIHHGILSNTHRQAARVNLSFDFPFYGHFLREITVATGGFIYTG
 EVVHRMLTATQYIAPLMANFDPSVSRNSTVRYFDNGTALVVQWDHVHLQDNYNLGSFTFQAT
 LLMDGRIIFGYKEIPVLVTQISSTNHPVKVGLSDAFVVVHRIQQIPNVRRRTIYEYHRVELQ
 MSKITNISAVEMTPLPTCLQFNRCGPCVSSQIGFNC SWCSKLQRCSSGFDRHRQDWVDSGCP
 EESKEKMCENTEPVETSSRTTTTVGATTTQFRVLTTRRAVTSQFPTSLPTEDDTKIALHLK
 DNGASTDDSAAEKKGGTLHAGLIIGILILVLIVATAILVTVYMYHHPTSAASIFFIERRPSR
 WPAMKFRRGSGHPAYAEVEPVGEKEGFIVSEQC

Important features of the protein:**Transmembrane domain:**

amino acids 454-478

N-glycosylation sites.

amino acids 103-107, 160-164, 213-217, 221-225, 316-320, 345-349

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 297-301, 492-496, 503-507

N-myristoylation sites.

amino acids 42-48, 100-106, 147-153, 279-285, 397-403, 450-456,
 455-461

67/75

FIGURE 67A

GCAGCCCTAGCAGGG**ATG**GACATGATGCTGTTGGTGCAGGGTGCTTGTTCGACACCAGTG
GCTGGCGGCGGTGCTCCTCAGCCTGTGCTGCCTGCTACCCTCCTGCCTCCCGGCTGGACAGA
GTGTGGACTTCCCCTGGGCGGCCGTGGACAACATGATGGTCAGAAAAGGGGACACGGCGGTG
CTTAGGTGTTATTTGGAAGATGGAGCTTCAAAGGGTGCCTGGCTGAACCGGTCAAGTATTAT
TTTTGCGGGAGGTGATAAGTGGTCAGTGGATCCTCGAGTTTCAATTTCAACATTGAATAAAA
GGGACTACAGCCTCCAGATACAGAATGTAGATGTGACAGATGATGGCCCATACACGTGTTCT
GTTTCAGACTCAACATACACCCAGAACAATGCAGGTGCATCTAACTGTGCAAGTTCCTCCTAA
GATATATGACATCTCAAATGATATGACCGTCAATGAAGGAACCAACGTCACTCTTACTTGTT
TGGCCACTGGGAAACCAGAGCCTTCCATTTCTTGGCGACACATCTCCCCATCAGCAAAACCA
TTTGAAAATGGACAATATTTGGACATTTATGGAATTACAAGGGACCAGGCTGGGAATATGA
ATGCAGTGCAGAAAATGATGTGTCTCCAGATGTGAGGAAAGTAAAGTTGTTGTCAACT
TTGCTCCTACTATTTCAGGAAATTAATCTGGCACCGTGACCCCGGACGCAGTGGCCTGATA
AGATGTGAAGGTGCAGGTGTGCCCTCCAGCCTTTGAATGGTACAAAGGAGAGAAGAAGCT
CTTCAATGGCCAACAAGGAATTATTATTCAAATTTTAGCACAGATCCATTCTCACTGTTA
CCAACGTGACACAGGAGCACTTCGGCAATTATACTTGTGTGGCTGCCAACAAGCTAGGCACA
ACCAATGCGAGCCTGCCTCTTAACCCCTCCAAGTACAGCCCAGTATGGAATTACCGGGAGCGC
TGATGTTCTTTTTCTCCTGCTGGTACCTTGTGTTGACACTGTCCTCTTTCACCAGCATATTCT
ACCTGAAGAATGCCATTCTACAAT**TAA**ATTCAAAGACCCATAAAAGGCTTTTAAGGATTCTCT
GAAAGTGCTGATGGCTGGATCCAATCTGGTACAGTTTGTAAAAGCAGCGTGGGATATAATC
AGCAGTGCTTACATGGGGATGATCGCCTTCTGTAGAATTGCTCATTATGTAAATACTTTAAT
TCTACTCTTTTTTGTATTAGCTACATTACCTTGTGAAGCAGTACACATTGTCCTTTTTTTAAG
ACGTGAAAGCTCTGAAATTACTTTTAGAGGATATTAATTGTGATTTTCATGTTTGTAATCTAC
AACTTTTCAAAGCATTTCAGTCATGGTCTGCTAGGTTGCAGGCTGTAGTTTACAAAAACGAA
TATTGCAGTGAATATGTGATTCTTTAAGGCTGCAATACAAGCATTTCAGTTCCCTGTTTCAAT
AAGAGTCAATCCACATTTACAAAGATGCATTTTTTTCTTTTTTGATAAAAAAGCAAATAATA
TTGCCTTCAGATTATTTCTTCAAATATAACACATATCTAGATTTTTCTGCTCGCATGATAT
TCAGGTTTCAGGAATGAGCCTTGTAATATAACTGGCTGTGCAGCTCTGCTTCTTTCTCTGT
AAGTTCAGCATGGGTGTGCCTTCATACAATAATATTTTTCTCTTTGTCTCCAACATAATAA
AATGTTTTGCTAAATCTTACAATTTGAAAGTAAAAATAAACCAGAGTGATCAAGTTAAACCA
TACACTATCTCTAAGTAACGAAGGAGCTATTGGACTGTAAAAATCTCTTCCTGCCTGACAA
TGGGGTTTGAGAATTTTGCCCCACACTAACTCAGTTCTTGTGATGAGAGACAATTTAATAAC
AGTATAGTAAATATACCATATGATTTCTTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGA
AATCATTCCCTTTAAATGACAGCACAGTCCACTCAAAGGATTGCCTAGCAATACAGCATCT
TTTCCTTTCACTAGTCCAAGCCAAAAATTTTAAGATGATTTGTCAGAAAGGGCACAAAGTCC
TATCACCTAATATTACAAGAGTTGGTAAGCGCTCATCATTAAATTTTATTTTGTGGCAGCTAA
GTTAGTATGACAGAGGCAGTGCTCCTGTGGACAGGAGCATTTTGCATATTTTCCATCTGAAA
GTATCACTCAGTTGATAGTCTGGAATGCATGTTATATATTTTAAACTTCCAAAATATATTA
TAACAAACATTCTATATCGGTATGTAGCAGACCAATCTCTAAAATAGCTAATTCTTCAATAA
AATCTTTCTATATAGCCATTTTCAGTGCAAACAAGTAAAATCAAAAAAGACCATCCTTTATTT
TTCTTACATGATATATGTAAGATGCGATCAAATAAAGACAAAACACCAGTGATGAGAATAT
CTTAAGATAAGTAATTATCAAATTATTGTGAATGTTAAATTATTTCTACTATAAAGAAGCAA
AACTACATTTTTTGAAGGAAAATGCTGTTACTCTAACATTAATTTACAGGAATAGTTTGATGG
TTTCACTCTTTACTAAAGAAAGGCCATCACCTTGAAAGCCATTTTACAGGTTTGATGAAGTT
ACCAATTTTCAGTACACCTAAATTTCTACAAATAGTCCCCTTTTACAAGTTGTAACAACAAAG
ACCCTATAATAAAATTAGATACAAGAAATTTTGCAGTGGTTATACATATTTGAGATATCTAG
TATGTTGCCCTAGCAGGGATGGCTTAAAACTGTGATTTTTTTCTTCAAGTAAACTTAGT
CCCAAAGTACATCATAAATCAATTTTAATTAGAAAAATGAATCTTAAATGAGGGGACATAAG
TATACTCTTTCCACAAAATGGCAATAATAAGGCATAAAGCTAGTAAATCTACTAACTGTAAT
AAATGTATGACATTATTTTGATTGATACATTAAAAAAGAGTTTTTAGAACAAATATGGCATT
TAACTTTATTATTTATTTGCTTTTAAGAAATATTCTTTGTGGAATTGTTGAATAAACTATAA
AATATTATTTTGATTGACAGCTTTAAAGTGGCACACTCCATAATAATCTACTTACTAGAAAT

FIGURE 67B

AGTGGTGCTACCACAAAAAATGTTAACCATCAGTACCATTGTTTGGGAGAAAGAAACAGATC
AAGAATGCATATTATTCAGTGACCGCTTTCCTAGAGTTAAAATACCTCCTCTTTGTAAGGTT
TGTAAGGTAAATTGAGGTATAAACTATGGATGAACCAAATAATTAGTTCAAAGTGTTGTCATG
ATTCCAAATTTGTGGAGTCTGGTGTTTTTACCATAGAATGTGACAGAAGTACAGTCATAGCT
CAGTAGCTATATGTATTTGCCTTTATGTTAGAAGAGACTTTCCTGAGTGACATTTTTAAATA
GAGGAGGTATTCATGATGTTTTTCTGTATCAGAGCAGCATTCCTAGTCCTTAGGCCCTCGGA
CAGAGTGAAATCATGAGTATTTATGAGTTCAATATTGTCAAATAAGGCTACAGTATTTGCTT
TTTTGTGTGAATGTATTGCATATAATGTTCAAGTAGATGATTTTACATTTATGGACATATAA
AATGTCTGATTACCCCATTTTATCAGTCCTGACTGTACAAGATTGTTGCAATTTTCAGAATAG
CAGTTTTATAAATTGATTTATCTTTTAATCTATAACAATTTGTGTTAGCTGTTTCATTTTCAGG
ANTATATTTTCTACAAGTTCCACTTGTGGGACTCCTTTTGTGCCCCCTATTTTTTTTTTAAAG
AAGGAAGAAAGAAAAATAAGTAGCAGTTTTAAAAATGAGAATGGAGAGAAAAAGAAAAAGAATG
AAAAGGAAAGGCAGTAAAGAGGGGAAAAAAAAGGAAGGATGGAAGGAATGAAGGAAGGAAGGG
AGGAAGGGGAGAAGGTAGGAAGAAAGAAAGGATGAGAGGGGAAGGAAGAATCAGAGTATTAGG
GTAGTTAACTTACACATTTGCATTCTTAGTTTAACTGCAAGTGGTGTAAGTATGTTTTTCAA
TGATCGCATTTGAAACATAAGTCCTATTATACCATTAAAGTTCCTATTATGCAGCAATTATAT
AATAAAAAGTACTGCCCAAGTTATAGTAATGTGGGTGTTTTTGTGAGACACTAAAAGATTTGAG
AGGGAGAATTTCAAACCTTAAAGCCACTTTTGGGGGGTTTATAACTTAACTGAAAAATTAATG
CTTCATCATAACATTTAAGCTATATCTAGAAAGTAGACTGGAGAACTGAGAAAATTACCCAG
GTAATTCAGGGAAAAAAAATATATATATATATAAATACCCCTACATTTGAAGTCAGAAA
ACTCTGAAAAACTGAATTATCAAAGTCAATCATCTATAATGATCAAATTTACTGAACAATTG
TTAATTTATCCATTGTGCTTAGCTTTGTGACACAGCCAAAAGTTACCTATTTAATCTTTTCA
ATAAAAATTGTTTTTTGAAATCCAGAAATGATTTAAAAAGAGGTCAGGTTTTTAACTATTTA
TTGAAGTATGTGGATGTACAGTATTTCAATAGATATGAATATGAATAAATGGTATGCCTTAA
GATTCTTTGAATATGTATTTACTTTAAAGACTGGAAAAAGCTCTTCCTGTCTTTTAGTAAAA
CATCCATATTTCATAACCTGATGTAAAATATGTTGTACTGTTTCCAATAGGTGAATATAAAC
TCAGTTTATCAATTAAAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT

ATAGTAGAAGAATGTCTCTGAAATTACTGGATGAGTTTCAGTCATATACTTTACATGGGCACA
ATTTACATTTCAAGCTCCTTATCCTAGGCTAATTTTATATTATGTTAAATCACTTGTTTTTG
TTCTCACGGCTTCCTGCCTGCTATAGGCATAATTACGAGGAAGCAGAACTTCTCCAGAAGCA
AGCGCACATGCGTTCCAAAATAAGAGCAAATTCGCTCTAAACACAGGAAAAGACCTGAAGCT
TTAATTAAGGGGTACATCCAACCCCAAGAGCGCTTTTGTGGGCACTGATTGCTCCAGCTTCT
GCGTCACTGCGCGAGGGAAGAGGGAAGAGGATCCAGGCGTTAGACATGTATAGACACAAAAA
CAGCTGGAGATTGGGCTTAAAATACCCACCAAGCTCCAAAGAAGAGACCCAAGTCCCCAAAA
CATTGATTTCAAGGCTGCCAGGAAGGAAGAGCAGCAGCAGGGTGGGAGAGAAGCTCCAGTCA
GCCCACAAGATGCCATTGTCCCCCGGCCTCCTGCTGCTGCTGCTCTCCGGGGCCACGGCCAC
CGCTGCCCTGCCCTGGAGGGTGGCCCCACCGGCCGAGACAGCGAGCATATGCAGGAAGCGG
CAGGAATAAGGAAAAGCAGCCTCCTGACTTTCCTCGCTTGGTGGTTTGAGTGGACCTCCCAG
GCCAGTGCCGGGCCCCCTCATAGGAGAGGAAGCTCGGGAGGTGGCCAGGCGGCAGGAAGGCGC
ACCCCCCAGCAATCCGCGCGCCGGGACAGAATGCCCTGCAGGAAGCTTCTTCTGGAAGACCT
TCTCCTCCTGCAAAATAG

71/75

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA44175

><subunit 1 of 1, 155 aa, 1 stop

><MW: 17194, pI: 10.44, NX(S/T): 0

MYRHKNSWRLGLKYPPSSKEETQVPKTLISGLPGRKSSSRVGEKLQSAHKMPLSPGLLLLLLL
SGATATAALPLEGGPTGRDSEHMQEAGIRKSSLLTFLAWWFEWTSQASAGPLIGEEAREVA
RRQEGAPPQQSARRDRMPCRNFFWKTFSSCK

Important features of the protein:

Transmembrane domain:

amino acids 51-69

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 35-39, 92-96

N-myristoylation sites.

amino acids 64-70, 75-81, 90-96

Amidation site.

amino acids 33-37

72/75

FIGURE 71

GTCGTGTGCTTGGAGGAAGCCGCGGAACCCCCAGCGTCCGTCC**ATG**GCGTGGAGCCTTGGGA
GCTGGCTGGGTGGCTGCCTGCTGGTGTGAGCATTGGGAATGGTACCACCTCCCGAAAATGTC
AGAATGAATTCTGTTAATTTCAAGAACATTCTACAGTGGGAGTCACCTGCTTTTGCCAAAGG
GAACCTGACTTTTACAGCTCAGTACCTAAGTTATAGGATATTCCAAGATAAATGCATGAATA
CTACCTTGACGGAATGTGATTTCTCAAGTCTTTCCAAGTATGGTGACCACACCTTGAGAGTC
AGGGCTGAATTTGCAGATGAGCATTGAGCTGGGTAAACATCACCTTCTGTCTGTGGATGA
CACCATTATTGGACCCCCCTGGAATGCAAGTAGAAGTACTTGCTGATTCTTTACATATGCGTT
TCTTAGCCCCCTAAAATTGAGAATGAATACGAACTTGGACTATGAAGAATGTGTATAACTCA
TGGACTTATAATGTGCAATACTGGAAAAACGGTACTGATGAAAAGTTTCAAATTACTCCCCA
GTATGACTTTGAGGTCCCTCAGAAACCTGGAGCCATGGACAACCTTATTGTGTTCAAGTTCGAG
GGTTTCTTCCTGATCGGAACAAAGCTGGGGAATGGAGTGAGCCTGTCTGTGAGCAAACAACC
CATGACGAAACGGTCCCCCTCCTGGATGGTGGCCGTCATCCTCATGGCCTCGGTCTTCATGGT
CTGCCTGGCACTCCTCGGCTGCTTCTCCTTGCTGTGGTGCCTTTACAAGAAGACAAAGTACG
CCTTCTCCCCCTAGGAATTCTCTTCCACAGCACCTGAAAGAGTTTTTGGGCCATCCTCATCAT
AACACACTTCTGTTTTTCTCCTTTCCATTGTGCGATGAGAATGATGTTTTTGACAAGCTAAG
TGTCATTGCAGAAGACTCTGAGAGCGGCAAGCAGAATCCTGGTGACAGCTGCAGCCTCGGGA
CCCCGCTGGGCAGGGGCCCAAAGC**TAG**GCTCTGAGAAGGAAACACACTCGGCTGGGCACA
GTGACGTACTCCATCTCACATCTGCCTCAGTGAGGGATCAGGGCAGCAAACAAGGGCCAAGA
CCATCTGAGCCAGCCCCACATCTAGAACTCCAGACCTGGACTTAGCCACCAGAGAGCTACAT
TTTAAAGGCTGTCTTGGCAAAAATACTCCATTTGGGAACTCACTGCCTTATAAAGGCTTTCA
TGATGTTTTTCAGAAGTTGGCCACTGAGAGTGTAATTTTCAGCCTTTTATATCACTAAAATAA
GATCATGTTTTAATTGTGAGAAACAGGGCCGAGCACAGTGGCTCACGCCTGTAATACCAGCA
CCTTAGAGGTCGAGGCAGGCGGATCACTTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAATA
TGGTGAAACCCAGTCTCTACTAAAAATACAAAAATTAGCTAGGCATGATGGCGCATGCCTAT
AATCCCAGCTACTCGAGTGCCTGAGGCAGGAGAATTGCATGAACCCGGGAGGAGGAGGGA
GGTTGCAGTGAGCCGAGATAGCGGCACTGCACTCCAGCCTGGGTGACAAAGTGAGACTCCAT
CTCAAAAAAAAAAAAAAAAAAATTGTGAGAAACAGAAATACTTAAAATGAGGAATAAGAATGG
AGATGTTACATCTGGTAGATGTAACATTCTACCAGATTATGGATGGACTGATCTGAAAATCG
ACCTCAACTCAAGGGTGGTCAGCTCAATGCTACACAGAGCACGGACTTTTGGATTCTTTGCA
GTACTTTGAATTTATTTTTCTACCTATATATGTTTTATATGCTGCTGGTGCCTCATTAAGT
TTTACTCTGTGTTGC

amino acids 6-12, 316-322

FIGURE 73

CGAGCGCCAACCCGCTAGCGCCTGAATCCGGCGTGCTGCCCGCTCGCCGCCCCGCC**ATG**CCCC
GCGCAGCCCCGCTGCTCGCCGCGTTGACCGCGCTCCTCGCCGCGCCGCTGCTGGCGGAGAT
GCCCCGCCGGGCAAAATCGCGGTGGTTGGGGCTGGGATTGGGGGCTCTGCTGTGGCCCATTT
TCTCCAGCAGCACTTTGGACCTCGGGTGCAGATCGACGTGTACGAGAAGGGAACCGTGGGTG
GCCGCTTGCCACCATCTCAGTCAACAAGCAGCACTATGAGAGCGGGGCTGCCTCCTTCCAC
TCCCTGAGCCTGCACATGCAGGACTTCGTCAAGCTGCTGGGGCTGAGGCACCGGCGCGAGGT
GGTGGGCAGGAGCGCCATCTTCGGCGGGGAGCACTTCATGCTGGAGGAGACTGACTGGTACC
TGCTGAACCTCTTCCGCCTCTGGTGGCACTATGGCATCAGCTTCCTGAGGCTGCAGATGTGG
GTGGAGGAGGTTCATGGAGAAGTTCATGAGGATCTATAAGTACCAGGCCACGGCTATGCCTT
CTCGGGTGTGGAGGAGCTGCTCTACTCACTGGGGGAGTCCACCTTTGTTAACATGACCCAGC
ACTCTGTGGCTGAGTCCCTGCTGCAGGTGGGCGTCACGCAGCGCTTTATTGATGATGTCGTT
TCTGCTGTCTGCGGGCCAGCTATGGCCAGTCAGCAGCGATGCCCGCCTTTGCAGGAGCCAT
GTCACTAGCCGGGGCCCCAAGGCAGCCTGTGGTCTGTGGAAGGAGGCAATAAGCTGGTTTGT
CCGGTTTGTCTGAAGCTCACCAAGGCCAATGTGATCCATGCCACAGTGACCTCTGTGACCCTG
CACAGCACAGAGGGGAAAGCCCTGTACCAGGTGGCGTATGAGAATGAGGTAGGCAACAGCTC
TGACTTCTATGACATCGTGGTCATCGCCACCCCCCTGCACCTGGACAACAGCAGCAGCAACT
TAACCTTTGCAGGCTTCCACCCGCCCATTTGATGACGTGCAGGGCTCTTTCCAGCCCACCGTC
GTCTCCTTGGTCCACGGCTACCTCAACTCGTCCTACTTCGGTTTCCCAGACCCTAAGCTTTT
CCCCTTTGCCAACATCCTTACCACAGATTTCCCCAGCTTCTTCTGCACTCTGGACAACATCT
GCCCTGTCAACATCTCTGCCAGCTTCCGGCGAAAGCAGCCCCAGGAGGCAGCTGTTTGGCGA
GTCCAGTCCCCCAAGCCCCCTCTTTCGGACCCAGCTAAAGACCCTGTTCCGTTCCCTATTACTC
AGTGCAGACAGCTGAGTGGCAGGCCCATCCCCTCTATGGCTCCCGCCCCACGCTCCCGAGGT
TTGCACTCCATGACCAGCTCTTCTACCTCAATGCCCTGGAGTGGGCGGCCAGCTCCGTGGAG
GTGATGGCCGTGGCTGCCAAGAATGTGGCCTTGCTGGCTTACAACCGCTGGTACCAGGACCT
AGACAAGATTGATCAAAAAGATTTGATGCACAAGGTCAAGACTGAAGTGTGAGGGCTCTAGG
GAGAGCCTGGGAACTTTCATCCCCCACTGAAGATGGATCATCCCACAGCAGCCCAGGACTGA
ATAAGCCATGCTCGCCCCACCAGGCTTCTTTCTGACCCCTCATGTATCAAGCATCTCCAGGTG
ACCTACTGTCTGCCTATATTAAGGGTCCACACGGCGGCTGCTGCTTTTTTTTAAAGGGGAAA
GTAAGAAAAGAGAAGGAAATCCAAGCCAGTATATTTGTTTTATTTATTTTTTTTAAAGAAGAA
AAAAGTTCATCTTCACAAGGTGCTTCAGACTTGGTTTTCTTAGCTAGAAACCAGAAGACTACG
GGAGGGAATATAAGGCAGAGAACTATGAGTCTTATTTTATTACTGTTTTTCACTACCTACTC
CCACAATGGACAATCAATTGAGGCAACCTACAAGAAAACATTTACAACCAGATGGTTACAAA
TAAAGTAGAAGGGAAGATCAGAAAACCTAAGAAATGATCATAGCTCCTGGTTACTGTGGACT
TGATGGATTTGAAGTACCTAGTTTCAGAACTCCCTAGTCACCATCTCCAAGCCTGTCAACATC
ACTGCATATTGGAGGAGATGACTGTGGTAGGACCCAAGGAAGAGATGTGTGCCTGAATAGTC
GTCACCATATCTCCAAGCTTCTTGGCAACCAGTGGGAAAAGAAACATGCGAGGCTGTAGGAA
GAGGGAAGCTCTTCTTGGCACCTAGAGGAATTAGCCATTCTCTTCTTATGCAAAGATTGA
GGAATGCAACAATATAAAGAAGAGAAGTCCCCAGATGGTAGAGAGCAGTCATATCTTACCCC
TAGATGTTTCATCCCAGCAGAAGAAAGAAGGTGTTGGGGTAGGATTCTTCAGAGGTTAGC
CTGGTACTTTCTCATCAGACACTAGCTTGAAGTAAGAGGAGAATTATGCTTTTCTTTGCTTT
TTCTACAAACCCTTAAAAATCACTTGTTTTTAAAAAGAAAGTAAAGCCCTTTTCATTCAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 74

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA116510

><subunit 1 of 1, 494 aa, 1 stop

><MW: 54646, pI: 7.27, NX(S/T): 6

MARAAPLLAALTALLAAAAAGGDAPPGKIAVVGAGIGGSAVAHFLQQHFGPRVQIDVYEKGT
VGGRLATISV NKQHYESGAASFHSLSLHMQDFVKLLGLRHRREVVGSAIFGGEHFMLEETD
WYLLNLFRLWWHYGISFLRLQMWWVEEVMEKFMRIYKYQAHGYAFSGVEELLYSLGESTFVNM
TQHSVAESLLQVGVTQRFIDDVSAVLRASYGQSAAMPAFAGAMSLAGAQGSLWSVEGGNKL
VCSGLLKLTKANVIHATVTSVTLHSTEGKALYQVAYENEVGNSSDFYDIVVIATPLHLDNSS
SNLTFAGFHPPIDDVQGSFQPTVVSILVHGYNSSYFGFPDPKLFPEANILTTDFPSFFCTLD
NICPVNISASFRRKQPQEA AVWRVQSPKPLFRTQLKTLFRSYYSVQTAEWQAHPYGSRPTL
PRFALHDQLFYLNLEWAASSVEVMAVA AKNVALLAYNRWYQDLDKIDQKDLMHKVKTEL

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 185-189, 290-294, 308-312, 312-316, 342-346, 378-382

N-myristoylation sites.

amino acids 33-39, 35-41, 38-44, 61-67, 64-70, 218-224, 234-240,
237-243, 429-435